

***CNTNAP2*-Promotorvarianten als mögliche Risikofaktoren für Autismus-Spektrum-Störungen**

Vom Fachbereich Biologie der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines

DOCTOR RERUM NATURALIUM

genehmigte Dissertation von Dipl.-Biol. Marnie Kopp aus Groß-Gerau

1. Referent: Prof. Dr. Bodo Laube
2. Referentin: Prof. Dr. Dipl.-Theol. Christine M. Freitag
3. Referent: Prof. Dr. Ralf Galuske

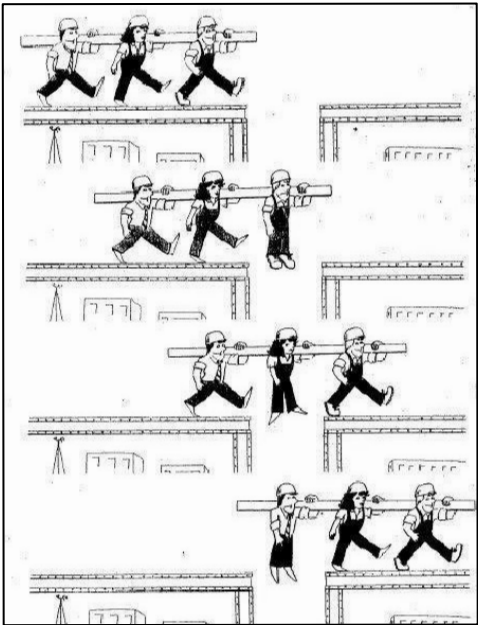
Tag der Einreichung: 28.10.2013

Tag der mündlichen Prüfung: 11.02.2014

Darmstadt, 2014

D 17

FÜR MEINE FAMILIE...



...UND MEINEN FRISÖR

Inhaltsverzeichnis

1	Zusammenfassung	1
2	Abstract	2
3	Einleitung	3
3.1	Autismus-Spektrum-Störungen	3
3.1.1	Ein Überblick	3
3.1.2	Symptomatik	5
3.1.3	Epidemiologie	7
3.1.4	Ätiologie	8
3.2	Genetische Studien bei Autismus-Spektrum-Störungen	8
3.2.1	Zwillings- und familienbasierte Studien	9
3.2.2	Kopplungsstudien	9
3.2.3	Assoziationsstudien	11
3.2.4	Genkopiepolymorphismen und chromosomale Aberrationen	12
3.2.5	Sequenzierungsstudien	13
3.3	Contactin-associated protein-like 2	14
3.3.1	Ein Überblick	14
3.3.2	<i>CNTNAP2</i> und Autismus-Spektrum-Störungen	15
3.3.3	Das <i>Cntnap2</i> -Mausmodell und Autismus-Spektrum-Störungen	18
3.3.4	Genexpressionsstudien und <i>CNTNAP2</i>	20
3.3.5	Promotorvarianten in der <i>CNTNAP2</i> -Promotorregion	20
3.4	Ziel der Studie	21
4	Material	22
4.1	Geräte	22
4.2	Sonstige Geräte und Verbrauchsmaterialien	23
4.3	Chemikalien	24
4.4	Puffer, Gele und Lösungen	25
4.4.1	DNA-Isolierung aus Wangenabstrichen	25
4.4.2	Polymerasekettenreaktion	26
4.4.3	Gelelektrophorese	26
4.4.4	Klonierung	26
4.4.5	Zellkultur	27
4.5	Probenmaterial	27

5	Methoden	29
5.1	Studiendesign	29
5.2	Isolierung der DNA aus dem Probenmaterial	30
5.2.1	Isolierung der DNA aus EDTA-Blut	30
5.2.2	Isolierung der DNA aus Wangenabstrichen	30
5.2.3	Konzentrationsbestimmung und Qualitätskontrolle der extrahierten DNA	31
5.3	Analyse der Detektionsstichprobe	32
5.3.1	Sequenzierung der <i>CNTNAP2</i> -Promotorregion	32
5.3.2	Auswertung der Sequenzierungsdaten	34
5.4	Analyse der Replikationsstichprobe	34
5.4.1	Genotypisierung der Replikationsstichprobe mittels RFLP	34
5.4.2	Genotypisierung der Replikationsstichprobe mittels Real-Time PCR	38
5.5	Funktionelle Analyse der Varianten	39
5.5.1	In-silico Analyse möglicher Transkriptionsfaktorbindestellen in der <i>CNTNAP2</i> -Promotorregion	39
5.5.2	Luciferase Assay	39
5.6	Statistische Auswertung	55
5.6.1	Deskriptive Statistik	55
5.6.2	Poweranalyse	55
5.6.3	Genetische Analyse	56
5.6.4	Auswertung Luciferase Assay	57
6	Ergebnisse	57
6.1	Beschreibung der Stichproben	57
6.2	Analyse der Detektionsstichprobe	58
6.2.1	Sequenzierung der <i>CNTNAP2</i> -Promotorregion	58
6.3	Analyse der Proben aus der Replikationsstichprobe und statistische Auswertung der Genotypisierungsdaten	63
6.3.1	Genotypisierung	64
6.3.2	Statistische Auswertung der Genotypisierungsdaten	64
6.4	Funktionelle Analyse der Varianten rs150447075, rs34712024 und rs71781329	67
6.4.1	In-silico Analyse	69
6.4.2	In-vitro Analyse	71
6.5	Zusammenfassung der Ergebnisse	77

7	Diskussion	78
7.1	Analysierte Stichprobe	78
7.2	Analyse der Detektionsstichprobe: Sequenzierung der <i>CNTNAP2</i> -Promotorregion	79
7.3	Analyse der Replikationsstichprobe: Genotypisierung und statistische Auswertung der Daten	81
7.4	Transfektionseffizienztest	83
7.5	Funktionelle Analyse der Varianten	85
7.6	Limitierung der Studie	89
7.7	Ausblick	90
8	Zusammenfassung und Schlussfolgerung	91
9	Literatur	92
10	Anhang	103
10.1	In dieser Studie verwendete Datenbanken	103
10.2	Abkürzungsverzeichnis	103
10.3	Sequenz der <i>CNTNAP2</i> -Promotorregion	107
11	Danksagung	108
12	Ehrenwörtliche Erklärung	110
13	Lebenslauf	111

1 Zusammenfassung

Autismus-Spektrum-Störungen (ASS) sind neuronale Entwicklungsstörungen des Kindesalters mit einer Erblichkeitsrate von 70-90%. Die Hauptsymptome sind Beeinträchtigungen der sozialen Interaktion und Kommunikation, abberante Sprachentwicklung, sowie repetitives und stereotypes Verhalten. Molekulargenetische Studien zeigten, dass *CNTNAP2* (contactin-associated protein-like 2) als Kandidatengen für ASS in Frage kommt. Dieses Gen liegt in einer der am besten replizierten Kopplungsregionen (7q35) für ASS und für mehrere Varianten des Gens konnte bereits eine Assoziation mit ASS nachgewiesen werden. Zudem wird vermutet, dass die Expressionslevel des Gens das ASS Risiko beeinflussen können. In der Region 1000bp vor dem Genstart sind 9 Einzel-Nukleotid-Polymorphismen (SNPs) und ein Short Tandem Repeat (STR) in beschrieben, welche trotz der Tatsache, dass Promotorvarianten die Genexpression beeinflussen können, bisher nicht im Kontext von ASS untersucht wurden. Das Ziel dieser Studie war es daher, diese und mögliche neue Varianten innerhalb des *CNTNAP2*-Promotors zu detektieren, die häufigsten Varianten auf eine Assoziation mit ASS zu testen und deren funktionellen Einfluss zu charakterisieren. Hierfür wurde die *CNTNAP2*-Promotorregion bei 667 Familienmitgliedern von 236 ASS-Familien direkt sequenziert und die häufigsten Varianten in 356 weiteren Familien (1102 Familienmitglieder) mittels RFLP und Real-Time PCR genotypisiert. Beide Stichproben wurden auf Assoziation mit ASS und Sprachentwicklung getestet. Die funktionelle Analyse der Varianten erfolgte in der fibroblastoiden Zelllinie HEK293T und während der Differenzierung der neuroblastoiden Zelllinie SH-SY5Y.

Durch die Sequenzierung konnten drei bekannte (rs150447075, rs34712024, rs71781329) sowie sechs neue Varianten nachgewiesen werden. Da die neuen Varianten jedoch von den Eltern an die Kinder vererbt wurden, bzw. gar nicht in den betroffenen Kindern vorkamen, wurden diese nicht weiter untersucht. Interessanterweise konnte für rs34712024G im Vergleich zu rs34712024A ein protektiver Effekt auf das ASS-Risiko sowie ein negativer Trend hinsichtlich der Assoziation mit Sprachentwicklung nachgewiesen werden. Auch die Träger von rs71781329GCG[7] zeigten eine verzögerte Sprachentwicklung im Vergleich zu GCG[6] bzw. GCG[8]. In-silico konnte auf funktioneller Ebene gezeigt werden, dass alle Varianten grundsätzlich Transkriptionsfaktorbindestellen ändern könnten. In-vitro zeigten die Minorallele einen Zell- und Differenzierungsstatus spezifischen Effekt auf die Promotoraktivität, wobei nur Varianten, die Bindestellen von hirnspezifischen Transkriptionsfaktoren ändern (rs34712024, rs71781329) auch einen Effekt auf das ASS Risiko und/oder die Sprachentwicklung hatten. Die hier gezeigten Ergebnisse lassen vermuten, dass die *CNTNAP2*-Promotorvarianten einen pleiotropen Effekt auf das ASS Risiko und die Sprachentwicklung ausüben. Der zelluläre Kontext spielt hierbei vermutlich durch die Aktivierung verschiedener Transkriptionsfaktoren, vor allem während der frühen Phase der neuronalen Entwicklung eine modulierende Rolle.

2 Abstract

Autism spectrum disorders (ASS) are predominantly genetically determined with an onset in early childhood and a high heritability of 70-90%. The core symptoms of this genetically complex neurodevelopmental disorder are impairments of social interaction and communication, stereotyped behaviour and interests as well as a marked deficit in language development. Previous genetic studies suggested the involvement of *CNTNAP2* (contactin-associated protein-like 2), a member of the Neurexin gene superfamily, as a candidate gene for ASS. This gene is located in one of the best replicated linkage regions (7q35) for ASS. Although, expression levels of *CNTNAP2* were proposed alter ASS risk and promoter variants are known to interfere with transcription factor binding sites, no study to date has focused on the genetic and functional analysis of the *CNTNAP2* promoter region. 1000 bp downstream of the gene 9 single nucleotide polymorphisms (SNPs) and a short tandem repeat (STR) are described. The aim of the present study was to screen for known and unknown promoter variants within the *CNTNAP2* promoter, to test the most common SNPs for association with ASS and to characterize their effect on the transcriptional efficiency of the promoter. Therefore, the *CNTNAP2* promoter region of 667 members of 236 German families each with one child diagnosed with ASS were directly sequenced using blood DNA. The three most frequent variants (rs150447075, rs34712024, rs71781329) were genotyped in additional 356 families (1102 members) using restriction fragment length polymorphisms (RFLP) and real-time PCR methods. Functional analyses of these variants were performed in fibroblastoid HEK293T cells and during neuronal differentiation of neuroblastoid SH-SY5Y cells.

Direct sequencing identified six novel previously undescribed, and three known variants (rs150447075, rs34712024, rs71781329). All novel variants were either parentally transmitted or not detected in any subjects with ASS diagnosis and thus not further investigated. Interestingly, in ASS subjects the minor G allele of rs34712024 was under-transmitted and showed a trend for association with language development. Carriers of rs71781329GCG[7] showed a delayed onset of speech. At the functional level, all variants were in-silico predicted to alter transcription factor binding sites. In-vitro, the minor alleles showed cell line specific and differentiation stage-dependent effects at the level of promoter activation and only variants that altered transcription factor binding sites related to neuronal activity (i.e. rs34712024 and rs71781329) showed an effect on ASS liability or language development. This result led to the conclusion that the effect of *CNTNAP2* promoter variants on ASS risk and language development is pleiotropic. This pleiotropy may be mediated by the specific cellular context and differential transcription factor activation during early stages of neuronal differentiation.

3 Einleitung

Spätestens durch den Film „Rainman“ wurde das Bild, das die Öffentlichkeit von Autisten hat, weitgehend festgelegt. So gelten Autisten als schusselige, aber liebenswerte Menschen, die gut Kopfrechnen können oder irgendeine andere besondere Fähigkeit haben. Zudem werden sie als zwanghafte Personen dargestellt, die sich leicht aufregen und auf gar keinen Fall in der Lage sind, sich selbst zu versorgen. Dass es sich hierbei jedoch um einen eher seltenen Subtypen (Asperger Syndrom mit Inselbegabung) eines wesentlich komplexeren Störungsbildes handelt, welches zudem schwer zu diagnostizieren ist und dessen Ätiologie bisher nicht geklärt werden konnte, ist den meisten nicht bekannt.

3.1 Autismus-Spektrum-Störungen

3.1.1 Ein Überblick

Der Begriff Autismus (*griechisch* αὐτός = selbst) wurde zum ersten Mal vor über 100 Jahren von Eugen Bleuler (Bleuler 1911) verwendet, um eines der vier Hauptsymptome der Schizophrenie, nämlich den durch den Rückzug in die eigene Gedankenwelt verursachten Realitätsverlust der Patienten, zu beschreiben. Die heutige Definition des Krankheitsbildes Autismus erfolgte erst viele Jahre später durch den Kinder- und Jugendpsychiater Leo Kanner (Kanner 1943) und den Kinderarzt und Heilpädagogen Hans Asperger (Asperger 1944). Beide berichteten in ihren Studien unabhängig von einander von Kindern mit einer scheinbar angeborenen Introversion, die Defizite in der Kommunikation aufweisen, kombiniert mit einem begrenzten Gebrauch von und Interesse an Mimiken und Gesten.

Da die Störung als sehr selten galt, wurde sie in den darauf folgenden Jahren nur in einem sehr geringen Umfang weiter erforscht. So wurden die autistischen Störungen auch erst 1980 in die dritte Auflage des Manuals psychischer Störungen der „American Psychiatric Association“ (DSM-III) aufgenommen. Der im Jahr 1981 eingeführte Begriff des „Asperger-Syndroms“ (Wing 1981) folgte sogar erst in den 1990er Jahren (ICD-10: 1992; DSM-IV: 1994). Derzeit werden die autistischen Störungen in den internationalen Klassifikationssystemen ICD-10 (World Health Organization 1992) und DSM-IV-TR (American Psychiatric Association 2000) in die Subgruppe der „tiefgreifenden Entwicklungsstörungen“ innerhalb der Gruppe der „Entwicklungsstörungen“ eingeordnet. Definiert werden die Entwicklungsstörungen hierin über eine im frühen Kleinkindalter stattfindende

Manifestation von Entwicklungseinschränkung oder -verzögerung in Funktionen, die mit der biologischen Reifung des Zentralnervensystems verknüpft sind und stetig, ohne Remissionen und Rezidive verlaufen. Die tiefgreifenden Entwicklungsstörungen sind zudem durch qualitative Abweichungen in unterschiedlichen Entwicklungsbereichen, insbesondere der sozialen Interaktion und der Kommunikation, gekennzeichnet. Auffällig ist hierbei auch das eingeschränkte, stereotype, sich wiederholende Repertoire an Interessen und Aktivitäten. Obwohl man beispielsweise für das Rett-Syndrom und die desintegrativen Störungen eine abweichende Ätiologie vermutet (Bölte 2009), werden verschiedene Diagnosen aufgrund der ähnlichen Symptomatik unter dem Begriff „tiefgreifende Entwicklungsstörungen“ zusammengefasst. Ist jedoch explizit von Autismus-Spektrum-Störungen (ASS) die Rede, umfasst dieser Begriff nur die folgenden, kategorialen Diagnosen: Frühkindlicher Autismus, Asperger-Syndrom und nicht näher spezifizierte pervasive Entwicklungsstörungen (PDD-NOS; nach DSM-IV-TR). Grundsätzlich kann zwar abhängig vom Krankheitsbeginn, Schweregrad und Intelligenzquotienten zwischen diesen drei Diagnosen unterschieden werden. Da die Diagnosegrenzen jedoch fließend in einander übergehen und subklinische Formen autistischen Verhaltens auch in der Allgemeinbevölkerung, insbesondere bei Verwandten ersten Grades der Patienten wieder zu finden sind (Bailey et al. 1998), gestaltet sich die Unterscheidung zwischen den einzelnen Diagnosen eher schwierig. Es wird daher versucht dem heterogenen Erscheinungsbild der autistischen Störungen durch eine quantitative Beschreibung gerecht zu werden und versteht diese als ein sich ähnelndes, nicht kategorial abgrenzbares Kontinuum (Bölte 2009), das an einem Ende bis in den Normbereich reicht. Es hat sich zudem durchgesetzt, bei der Diagnose des klassischen oder auch frühkindlichen Autismus zwischen High-functioning-Autismus (HFA; $IQ > 70$) und Low-functioning-Autismus (LFA; $IQ \leq 70$) zu unterscheiden. Die Diagnose HFA bezieht sich hierbei auf das hohe Funktionsniveau der Patienten, die jedoch aufgrund ihrer Auffälligkeiten in der Sprachentwicklung die Diagnose eines Asperger-Syndroms nicht erhalten haben. Innerhalb der dimensional Klassifikation gilt der frühkindliche Autismus, aufgrund der starken Symptome, als schwerste und das Asperger Syndrom, welches an der Grenze zum klinisch nicht relevanten Bereich steht, als mildeste Form der ASS. Dazwischen sind alle Diagnosen des atypischen Autismus (PDD-NOS) anzuordnen (Abbildung 1). Als Goldstandard für die Diagnosestellung gelten die Beobachtungsskala für Autistische Störungen (ADOS-G; Lord et al. 2000) und das Diagnostische Interview für Autismus - Revidiert (ADI-R; Lord et al. 1994), in dem die Eltern oder mit der Entwicklung und der aktuellen Symptomatik des Probanden vertrauten Bezugspersonen zur frühkindlichen Entwicklung, zu Spracherwerb und möglichem Verlust

von sprachlichen Fertigkeiten, verbalen und non-verbalen kommunikativen Fähigkeiten, Spiel- und sozialem Interaktionsverhalten, stereotypen Interessen und Aktivitäten sowie komorbiden Symptomen (Aggression, Selbstverletzung, Epilepsie) befragt werden.

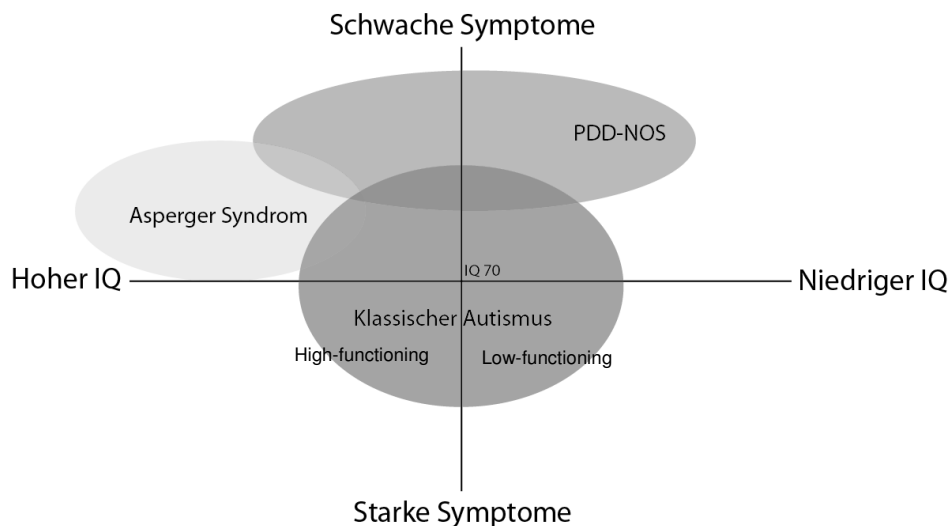


Abbildung 1: Subtypen – Asperger Syndrom, klassischer oder frühkindlicher Autismus und nicht anderweitig spezifizierte pervasive Entwicklungsstörungen (PDD-NOS) – variieren innerhalb der Symptome und Schweregrade. Bei dem klassischen Autismus wird zudem zwischen dem an den Normbereich grenzenden High-functioning- und dem Low-functioning-Autismus unterschieden (nach: Chiocchetti 2012).

3.1.2 Symptomatik

Die bei der Diagnosestellung beobachteten Symptome des frühkindlichen Autismus, des Asperger-Syndroms und der nicht näher spezifizierten pervasiven Entwicklungsstörungen (PDD-NOS) können von Patient zu Patient sehr von einander abweichen und sind auch aufgrund der individuellen Ausprägung der Störung sehr vielfältig. Auch sind nicht bei allen Diagnosen die soziale Interaktion, die Kommunikation und das Verhalten in gleicher Weise betroffen. Bei Kindern mit einem frühkindlichen Autismus liegen zwar bereits vor dem dritten Lebensjahr Defizite in allen drei Bereichen vor, Patienten mit Asperger-Syndrom sind bei Diagnosestellung jedoch bereits älter und weisen keine besonderen Auffälligkeiten in der Sprachentwicklung auf. Die Diagnose PDD-NOS wird hingegen nur dann gestellt, wenn das Kind die Symptome erst nach dem dritten Lebensjahr in mindestens einem der Bereiche entwickelt. Zudem können die Symptome von geradezu unauffälligen Verhaltensproblemen bis hin zu einer schweren geistigen Behinderung variieren (World Health Organization 1992, American Psychiatric Association 2000):

Soziale Interaktion

Die soziale Interaktion kann bei Kindern mit ASS stark beeinträchtigt sein. Sie zeigen häufig eine starke Objektbezogenheit, die meist auf eine bestimmte Gegenstandsart begrenzt ist. Auch in der Nutzung und dem Verständnis sozialer und emotionaler Kommunikationsmittel, wie Mimik und Gestik, sind autistische Kinder stark eingeschränkt. Es ist ihnen nicht möglich Gefühle oder Interessen mit anderen Personen zu teilen, oder von ihrem Gegenüber wahrzunehmen, wodurch sich die Beziehung zu ihren Mitmenschen auf den Austausch von Informationen ohne sozial-kommunikative Komponenten begrenzt.

Kommunikation

Die Kommunikation mit Autisten gestaltet sich, wenn überhaupt möglich, als sehr schwierig, da sie häufig in ihren verbalen und nonverbalen Fähigkeiten eingeschränkt sind. Neben Sprachentwicklungsverzögerungen kann es hierbei zum Sprachentwicklungsabbruch oder dem völligen Ausbleiben der Sprache kommen, wobei es den Betroffenen nicht möglich ist, die Defizite in der verbalen Sprache durch nonverbale, wie Mimik oder Gestik, auszugleichen. Sprichwörter, Redewendungen und Ironie können betroffene Kinder meistens nicht richtig deuten, da sie das Gesagte aufgrund von mangelnder Kreativität und Phantasie wörtlich auffassen. Weitere Auffälligkeiten liegen auch in der Intonation und der allgemeinen Betonung, sowie in einem idiosynkratischen, pedantischen Sprachstil.

Verhalten

Die autistmustypischen Störungen im Verhalten, die stark durch Routine und festgelegte Abläufe bestimmt werden, äußern sich vor allem in sich wiederholenden („repetitiven“) und stereotypen Verhaltensmuster, Interessen und Aktivitäten. Abweichungen können die Betroffenen extrem verunsichern oder ängstigen, wodurch es nötig wird, Tagesabläufe, Regeln und Zeitabsprachen genau einzuhalten. Als sehr sonderbar wird von Außenstehenden vor allem das stereotype, motorische Verhalten angesehen. Hierzu gehören vor allem Jaktationen wie heftige Schaukelbewegungen des Kopfes und des Körpers, Hand- und Fingermanierismen, aber auch extreme Verdrehungen des Körpers. Nicht selten wird auch ein selbst verletzendes Verhalten, wie zum Beispiel intensives Kratzen oder Beißen bei den Betroffenen beobachtet.

3.1.3 Epidemiologie

Basierend auf verschiedenen epidemiologischen Studien wird die Prävalenz für Autismus-Spektrum-Störungen (ASS) meist in einem Bereich von 0,9 bis 1,1% (Überblick in: Freitag 2012) angegeben, wobei die Angaben in Abhängigkeit der eingesetzten Diagnoseinstrumente und der Ausgangsstichprobe variieren können. So schwankt die Prävalenzrate im europäischen Raum zwischen 30-116/10.000, mit einem durchschnittlichen Wert von 62/10.000. Daher ist anzunehmen, dass von 160 Kindern mindestens eines Symptome des autistischen Spektrums aufweist (Elsabbagh et al. 2012). Autismus-Spektrum-Störungen gehören mittlerweile zu den am häufigsten diagnostizierten kindlichen Entwicklungsstörungen, für die in den letzten Jahren ein allgemeiner Anstieg in der Häufigkeit verzeichnet werden konnte (Fombonne 2003, 2005), wobei angenommen wird, dass dieser Anstieg vor allem dem heutigen Bekanntheitsgrad von ASS und der deutlich verbesserten Diagnostik zuzuschreiben ist, dass also nicht die Zahl der Erkrankungen, sondern vielmehr die Zahl der diagnostizierten Fälle angestiegen ist (Wing und Potter 2002).

Was jedoch aus allen Studien hervorgeht ist, dass Jungs häufiger betroffen sind als Mädchen, wobei das Geschlechterverhältnis durchschnittlich bei 3-4:1 liegt (Volkmar et al. 2004). Bei komorbider geistiger Behinderung sinkt das Verhältnis jedoch auf 2:1 (Fombonne 2002) und kann bei einem hohen Intelligenzbereich auf 11:1 und mehr ansteigen (Gillberg et al. 2006). Diese mit dem Intelligenzniveau zusammenhängende Varianz des Geschlechterverhältnisses liegt darin begründet, dass Mädchen mit ASS meist eine schwere geistige Behinderung aufweisen (Lord und Schopler 1985). Aus einem aktuellen Review (Kirkovski et al. 2013) geht jedoch hervor, dass Frauen möglicherweise nicht korrekt diagnostiziert werden. Bei ihnen liegt der entsprechende Phänotyp nämlich meist so verändert vor, dass er nur schwer zu diagnostizieren ist, bzw. nicht adäquat durch die entsprechenden Diagnoseinstrumente abgedeckt wird. Hinzu kommt, dass Frauen mit ASS häufig eine Strategie entwickeln, um bestimmte Alltagssituationen zu bewältigen, wobei sie ihre Symptome einfach überspielen. Dadurch ist die weibliche Form von ASS noch weitgehend unbekannt, was wiederum häufig zu Fehlinterpretationen und Missverständnissen führt bzw. durch die zur Verfügung stehenden Diagnoseinstrumente unzureichend erfasst wird.

3.1.4 Ätiologie

Bei etwa 10-15% der ASS-Patienten tritt die autistische Störung begleitend zu einer gut charakterisierten Grunderkrankung mit spezifischer Ätiologie auf. Dazu gehören unter anderem die Tuberöse-Sklerose, das Fragile-X-Syndrom, das Smith–Lemli–Optiz Syndrom, oder das Timothy-Syndrom (Abrahams und Geschwind 2008). Bei etwa 30% der Kinder mit ASS tritt zusätzlich eine Form der Epilepsie auf (Tuchman und Rapin 2002), aufgrund dessen sie als häufigste Begleiterkrankung von ASS gilt.

Hinsichtlich der Ätiologie von ASS ist man bereits den unterschiedlichsten Hypothesen nachgegangen. Ursprünglich wurde vermutet, dass vor allem das Elternhaus entscheidend an der Entstehung von ASS beteiligt ist und überwiegend Kinder aus gut situierten Elternhäusern betroffen sind (Kanner 1943). Diese Vermutung konnte jedoch durch verschiedene epidemiologische Studien widerlegt werden (Schopler et al. 1979; Fombonne 2003). Auch erziehungsunabhängige Umweltfaktoren wurden bereits häufiger als Ursache für ASS diskutiert. So sind neben zu hohen Quecksilber- und sonstigen Umweltbelastungen (Mutter et al. 2005; Windham et al. 2006; Doja und Roberts 2006) auch das Alter der Eltern, sowohl der Mütter als auch der Väter, als mögliche Faktoren für die Entstehung von ASS in den Fokus geraten (Reichenberg et al. 2006; Weintraub 2011). Deren tatsächlicher Einfluss an der Entstehung des Krankheitsbildes konnte jedoch bisher nicht belegt werden.

Ein großes Augenmerk liegt derzeit auf der Untersuchung verschiedenster genetischer Risikofaktoren, aufgrund deren Interaktion der autistustypische Phänotyp in einem komplexen, multifaktoriellen Vererbungsmodell entsteht (Holtermann et al. 2006; Klauck 2006). Es wurde eine Beteiligung von 3-4 bis zu 100 betroffenen Genen diskutiert (Pickles et al. 1995; Pritchard 2001). Die eigentliche Ursache der Autismus-Spektrum-Störungen ist jedoch trotz der bisherigen Forschungsansätze bis heute weitgehend unbekannt.

3.2 Genetische Studien bei Autismus-Spektrum-Störungen

Obwohl für ASS bereits eine Erblichkeitsrate von 70-90% (Überblick in: Freitag et al. 2010; Hallmayer et al. 2011) nachgewiesen wurde, konnte die Frage der Beteiligung verschiedenster Gene bzw. Genloci oder Genvarianten an der Entstehung von Autismus bisher nicht vollständig aufgeklärt werden. Da bei Autisten neuroanatomische Besonderheiten jedoch keine Seltenheit sind (z.B. Courchesne et al. 2007; Stanfield et al. 2008), geht man davon aus, dass durch die

vorliegenden Genvarianten in Komponenten essenzieller Stoffwechselwege, die an der Ausbildung unterschiedlichster Hirnstrukturen beteiligt sind, entscheidende Veränderungen hervorgerufen werden. Um einen Einblick in die autistischen Störungen und deren Ätiologie zu erhalten greift man daher auf die unterschiedlichsten Methoden zurück:

3.2.1 Zwillings- und familienbasierte Studien

Um die Rekurrenz, also das wieder auftreten von Erkrankungen innerhalb von betroffenen Familien und somit die Erblichkeit komplexer Krankheiten zu beurteilen, werden im Allgemeinen Studien mit monozygoten und dizygoten Zwillingen, bzw. familienbasierte Studien durchgeführt. Aus diesen Studien kann auch die Heritabilität, also der Grad der Aufklärung der phänotypischen Varianz durch additive genetische Effekte, berechnet werden. Basierend auf unabhängig von einander durchgeführten Zwillingstudien wird die Konkordanzrate für die Entstehung von autistischen Störungen in der Literatur für monozygote Zwillinge mit 36-91% angegeben. Für dizygote Zwillinge liegt die Konkordanzrate hingegen bei 0-30% (Folstein und Rutter 1977; Steffenburg et al. 1989; Le Couteur et al. 1996; Bailey et al. 1998; Lichtenstein et al. 2010; Hallmayer et al. 2011). Man geht daher davon aus, dass ASS eine Heritabilität von 70-90% zeigen und somit überwiegend genetisch verursacht sind (Überblick in: Freitag et al. 2010; Hallmayer et al. 2011). Untermuert wird dies durch die Ergebnisse familienbasierter Studien, die bei Verwandten ersten Grades ebenfalls subklinische autistische Verhaltensweisen in sozialen, kognitiven und verbalen Bereichen nachweisen konnten (Bölte et al. 2007; Sasson et al. 2013).

3.2.2 Kopplungsstudien

Kopplungsstudien dienen dem Nachweis einer Kosegregation eines polymorphen Markers mit einer bestimmten Krankheit. Sie beruhen darauf, dass die Wahrscheinlichkeit einer stattfindenden Rekombination (Crossing-over) zwischen zwei Chromosomenabschnitten in der Meiose umso geringer wird, je näher sie zueinander liegen. So kann die Kopplung zwischen einer Erkrankung und einem bekannten Marker einen Hinweis auf die Lokalisation eines bis dato unbekannten krankheitsverursachenden Genort geben (Campbell 1997). Die Wahrscheinlichkeit, dass zwei Loci am selben Chromosom nah beieinander liegen, wird durch den Logarithmus der Wahrscheinlichkeit von Kopplung gegenüber Nichtkopplung (LOD-Score = Logarithm Of The Odds) angegeben, wobei ein LOD-Score von 3 bedeutet, dass eine tausendfach erhöhte Wahrscheinlichkeit (10^3) vorliegt, dass der Marker aufgrund einer Kopplung und nicht durch Zufall in Verbindung mit der Erkrankung

auftritt. Für das Studiendesign gibt es zum einen die Möglichkeit ein genomweites Screening oder ein so genanntes „Fine-mapping“ durchzuführen. Bei Letzterem werden Polymorphismen von Genen untersucht, deren Genprodukte bereits als pathophysiologisch relevant beschrieben worden sind.

In diesen Studien wird vor allem auf Multiplexfamilien zurückgegriffen, bei denen mehrere Familienmitglieder von der jeweiligen Erkrankung betroffen sind. So ist es möglich spezifische Kopplungsmuster nachweisen, durch die wiederum Chromosomenabschnitte identifiziert werden können, die nur an Patienten, nicht aber an gesunde Nachkommen weitergegeben wurden. Dadurch erhält man zwar einen Hinweis darauf, welche Regionen mit der Krankheit gekoppelt sind und welche Gene an der Krankheitsentstehung beteiligt sein könnten, der eigentliche krankheitsverursachende Mechanismus geht daraus jedoch nicht hervor.

Durch unabhängige Analysen konnte so bereits für die Regionen 2q21–33, 3q25–27, 3p25, 4q32, 6q14–21, 7q22, 7q31–36, 10p12–q11.1, 11p12–13, 17p11.2–q12 und 17q11–21 eine Kopplung mit möglichen krankheitsverursachenden Genloci für ASS nachgewiesen werden (Überblick in: Freitag et al. 2010). Die Gene, die innerhalb dieser gekoppelten Bereiche liegen kommen daher als Kandidatengene für ASS in Frage. Viele der Ergebnisse ließen sich häufig nicht reproduzieren, was die Autoren sowohl auf die genetische und phänotypische Heterogenität des autistischen Spektrums (Szatmari et al. 2007; Abrahams und Geschwind 2008), als auch auf das generelle Studiendesign zurückführen (Überblick in: Freitag et al. 2010). Nichtsdestotrotz gilt die um das Gen *CNTNAP2* (Contactin-associated protein-like 2) gelegene Region 7q22–q36 als die am besten replizierte Kopplungsregion bei ASS (Überblick in: Freitag et al. 2010).

In einer molekulargenetischen Studie, die 1998 vom „International Molecular Genetic Study of Autism Consortium“ durchgeführt wurde, konnte so bereits gezeigt werden, dass in 87 betroffenen Geschwisterpaaren für einen Marker der Region 7q32.1 ein LOD-Score von 2,53 vorliegt (International Molecular Genetic Study of Autism Consortium 1998). In der 152 betroffene Geschwisterpaare umfassenden Folgestudie konnte hingegen für einen Marker in der Region 7q22.1 ein LOD-Score von 3,20 nachgewiesen werden (International Molecular Genetic Study of Autism Consortium 2001). In einer weiteren Studie, bei der 160 Geschwisterpaare mit einer strikten Autismusdiagnose untersucht wurden, lag für einen Marker der Region 7q36.1 der LOD-Score bei 2,13 (Liu et al. 2001). 2002 konnte in einer Studie mit 99 Multiplexfamilien gezeigt werden, dass für einen Marker der Region 7q36 ein LOD-Score von 1,66 vorliegt (Shao et al. 2002) Durch diese und

andere bis heute durchgeführten Kopplungsstudien konnten somit für 7q22-36 immer wieder LOD-Scores > 1 aufgezeigt werden (Überblick in: Freitag et al. 2010). In dieser Region liegt unter anderem das Gen *CNTNAP2*, welches in dieser Studie genauer untersucht werden soll.

3.2.3 Assoziationsstudien

Durch die vollständige Sequenzierung des menschlichen Genoms (Human Genome Project; Venter et al. 2001), den Ergebnissen des internationalen HapMap-Projektes (<http://www.hapmap.org>; International HapMap Consortium 2005) und weiteren Forschungsprojekten wie dem internationalen 1000-Genome-Projekt sind derzeit fast 54 Millionen Referenznummern für genetische Varianten (RELEASE (27.06.2012): NCBI dbSNP Build 137 for Human; Stand: 28.05.2013) in der Datenbank vom „National Center for Biotechnology Information“ (dbSNP: <http://www.ncbi.nlm.nih.gov/SNP/>) erfasst und stehen der Erforschung ihres möglichen genetischen Beitrags bei der Entstehung der unterschiedlichsten Krankheitsbilder zur Verfügung. Basierend auf der Verteilung von Allelen eines Einzelnukleotid-Polymorphismus (SNPs) kann in Assoziationsstudien untersucht werden, ob bei Merkmalsträgern, bzw. Patienten ein Allel oder eine bestimmte Kombination von Allelen mehrerer SNPs (Haplotyp) häufiger nachgewiesen werden kann, als bei gesunden Kontrollen. Das eigentliche Ziel dieser Studien ist daher die gemeinsam mit der Erkrankung auftretenden spezifischen Ausprägungen der Gene nachzuweisen. Aufgrund von genomweit durchgeführten Assoziationsstudien (GWAS) konnte bereits eine lange Liste von ASS relevanten Genen und Varianten erstellt werden. Die meisten dieser Assoziationsbefunde konnten jedoch auch hier nicht repliziert werden, was wiederum auf eine Populationsstratifikation der Stichproben zurückzuführen sein könnte (Curran et al. 2011).

In dem bereits beschriebenen Kandidatengen *CNTNAP2* der Kopplungsregion 7q22-q36 konnte für mehrere SNPs eine Assoziation mit ASS nachgewiesen werden. So konnten beispielsweise Arking und Kollegen (2008) in ihrer Studie mit 72 Multiplexfamilien zeigen, dass für die *CNTNAP2*-Variante rs7794745T eine signifikante Assoziation mit ASS vorliegt ($p\text{-Wert} < 2,14 \times 10^{-5}$). Zur selben Zeit wurde an 304 Eltern-Kind-Trios mit ASS des AGRE-(Autism Genetic Resource Exchange)-Samples eine weitere Assoziationsstudie durchgeführt. Hierbei wurden in 7q35, welche auch als Kopplungsregion für Sprache gilt, 2758 SNPs genotypisiert. Es zeigte sich, dass für den Einzelnukleotidpolymorphismus rs2710102C eine signifikante Assoziation ($p\text{-Wert} = 0,028$) mit dem Sprachphänotyp in ASS vorliegt und diesen negativ beeinflusst (Alarcón et al. 2008). Die Assoziation

von diesen Varianten weist somit darauf hin, dass *CNTNAP2* eine Rolle bei der Entstehung des Krankheitsbildes spielt.

3.2.4 Genkopiepolymorphismen und chromosomale Aberrationen

Bei 3-5% der Patienten mit ASS können so genannte „Copy number variations“ (CNV) ursächlich für ASS sein. CNVs sind strukturelle Veränderungen des Erbgutes, bei denen eine abweichende Kopienzahl eines bestimmten DNA-Abschnittes innerhalb eines Genoms vorliegt. Die Analyse der jeweils vorliegenden Anzahl der Genkopien hat bei der Erforschung von ASS an Aufmerksamkeit gewonnen. Es konnte nämlich festgestellt werden, dass die Frequenz der nachgewiesenen *de novo*-CNVs bei Personen mit ASS deutlich erhöht ist (Sebat et al. 2007; Marshall et al. 2008). Zudem zeigte sich, dass bei Patienten mit ASS CNVs häufiger kodierende als nicht kodierende Bereiche betroffen sind und dass diesen Genen (z.B. *NRXN1*, *NLGN3/4X* und *SHANK3*) überdurchschnittlich häufig eine Beteiligung an ASS oder einer geistigen Behinderung nachgewiesen werden konnten (Pinto et al. 2010). In mehreren Fallstudien konnte auch ein Zusammenhang zwischen spezifischen Gen- oder Genabschnittsdeletionen innerhalb der Region 7q34–q36.2 mit geistigen Behinderungen, Sprachentwicklungsstörungen bzw. ASS hergestellt. Beispielsweise konnte in einem 9jährigen Mädchen eine 5,27 Mb große Deletion nachgewiesen werden [46,XX,del(7)(q36.1-q36.2)] (Caselli et al. 2008), die phänotypisch im Zusammenhang mit einer geistigen Behinderung, sowie ungewöhnlichen Gesichtszüge, einer Hypoplasie der Niere und einem Long QT Syndrom bei diesem Kind steht. Eben dieser Phänotyp konnte auch bereits in einer früheren Fallstudie beobachtet werden, in der bei dem untersuchten Zwillingsspaar eine umfangreichere Deletion in der gleichen Region vorlag [46,XX,del(7)(q34q36.2)] (Bisgaard et al. 2006). In einer weiteren Fallstudie konnte eine interstitielle Deletion eines 12 Mb großen Abschnittes in der chromosomalen Region 7q33q36 nachgewiesen werden. Diese Haploinsuffizienz, die auch das Gen *CNTNAP2* umfasste, führte bei der Patientin unter anderem zu einer geistigen Behinderung, ASS und einem fast vollständig fehlenden und stark beeinträchtigten Sprachvermögen (Rossi et al. 2008). Diese phänotypischen Beeinträchtigungen konnten auch bei einem weiteren Geschwisterpaar beobachtet werden, in denen eine 12,2 Mb Deletion in der chromosomalen Region 7q34–q36.2 nachgewiesen werden konnte (Sehested et al. 2010). In einer genomweiten CNV Studie mit ASS Familien konnte zudem ein maternal vererbter CNV in der *CNTNAP2*-Promotorregion eines betroffenen Kindes nachgewiesen werden, welcher bei den CNV Trägern (Mutter und Kind) im Vergleich zur Kontrollgruppe in einer

verringerten *CNTNAP2*-Expression resultierte (Nord et al. 2011). Somit wird eine mögliche Beteiligung des *CNTNAP2*-Gens an der Entstehung von ASS weiter bestärkt.

3.2.5 Sequenzierungsstudien

Die Sequenzierung dient zur Bestimmung der Nukleotid-Abfolge in einem DNA-Molekül und erlaubt so die Identifizierung neuer und möglicherweise krankheitsverursachender Genvarianten. Hierbei werden einerseits gezielt einzelne Kandidatengene meist mittels Sanger-Sequenzierung analysiert, und andererseits die gesamten kodierenden Bereiche des Genoms (Whole-Exome-Sequencing) oder der kompletten genomischen Region (Whole-Genome-Sequencing) mittels der neu entwickelten „Next-Generation-Sequencing“-Methoden analysiert.

In einer Studie zur Identifikation von krankheitsverursachenden Mutationen bei Patienten mit fokalen corticalen Dysplasie (FCD) konnte erstmals durch Sanger-Sequenzierung der Exons 9 bis 24 von *CNTNAP2* eine rezessive Mutation nachgewiesen werden, die das Risiko für ASS bei Patienten mit FCD erhöht (Strauss et al. 2006). FCD ist eine seltene Störung in der Entwicklung der Großhirnrinde, die nicht nur eine erhöhte Prävalenz für ASS zeigt sondern auch durch epileptische Anfälle, geistige Behinderungen, Hyperaktivität und eine Rückentwicklung der Sprache gekennzeichnet ist.

In einer weiteren Studie, in der die kodierenden Bereiche des *CNTNAP2*-Gens in 635 Patienten mit ASS und 942 Kontrollen Sanger-sequenziert wurden, konnten 27 stille Mutationen nachgewiesen werden, von denen 8 ausschließlich in Patienten vorlagen. Eine stille Variante in Exon 17, welche bei vier betroffenen Kindern vorkam, lag zudem in einer hoch konservierten Region (Bakkaloglu et al. 2008). Auch wenn stille Mutationen keinen Einfluss auf die Aminosäuresequenz haben, können dennoch Veränderungen in der Translationskinetik entstehen, was wiederum zur Fehlfaltung von Proteinen oder veränderter Proteindynamik und somit zum Auslösen von Krankheitssymptomen führen kann (Zhang et al. 2009).

In einer Next-Generation-Sequenzierungsstudie konnten durch eine Exomsequenzierung von 20 Kindern mit ASS und deren Eltern 21 de novo Mutationen nachgewiesen werden, von denen 11 die Struktur der daraus resultierenden Proteine verändern. Eine dieser strukturverändernden Varianten liegt unter anderem im *CNTNAP2*-Gen, welche im Zellmodell zu einer Überexpression des *CNTNAP2*-Proteins Caspr2 führt (O'Roak et al. 2011).

Die hier genannten Sequenzierungsstudien zeigen, dass auch seltene *CNTNAP2*-Varianten bzw. Mutationen das Risiko für ASS bzw. ASS typische Phänotypen wie zum Beispiel Sprachentwicklungsstörungen erhöhen. Es kann somit vermutet werden, dass sowohl häufige als auch seltene Varianten des *CNTNAP2*-Gens eine Rolle bei ASS spielen (vergl. auch Alarcón et al. 2008; Vernes et al. 2008; Newbury et al. 2011). Der krankheitsauslösende Mechanismus der einzelnen *CNTNAP2*-Varianten, sowie deren Einfluss auf die Funktion des Proteins bzw. die Genexpression konnte bisher nicht geklärt werden.

3.3 Contactin-associated protein-like 2

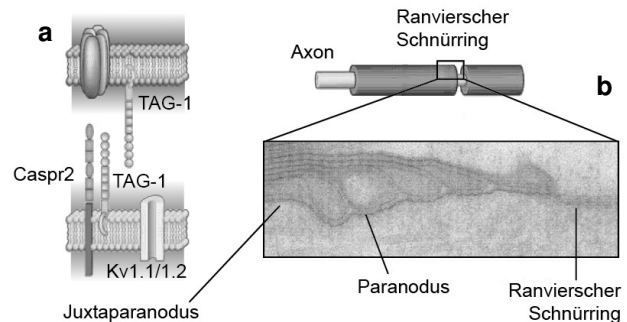
3.3.1 Ein Überblick

Die Ergebnisse aus den Kopplungs-, Assoziations- CNV- und Sequenzierungsstudien weisen alle auf eine Beteiligung von *CNTNAP2* an der Entstehung von ASS hin. Da Varianten des *CNTNAP2*-Gens bereits häufiger mit Krankheitsbildern in Verbindung gebracht werden konnten, die mit der neuronalen Entwicklung zusammenhängen, ist es wichtig die Funktionsweise und molekulare Aufgabe des Gens besser zu verstehen. Ein Großteil der Informationen, die uns bisher für das Gen vorliegen basieren auf Studien, die postnatal in den Gehirnen von Mäusen durchgeführt wurden (Überblick in: Peñagarikano und Geschwind 2012).

Das 25 Exons enthaltende Gen der chromosomalen Region 7q35–q36.1 erstreckt sich im Menschen über die Länge von 2,3 Mb (1,5 % des Chromosoms 7) und kodiert als eines der größten Gene im menschlichen Genom für das 180-kD schwere Transmembranprotein Caspr2 (Nakabayashi und Scherer 2001). 1999 wurde das *CNTNAP2*-Protein Caspr2 erstmals in Säugetieren als das Homolog des in *Drosophila melanogaster* vorkommenden Neurexin IV identifiziert, welches als neues Mitglied der Familie der Neurexine an der Interaktion zwischen Neuron und Glia am myelinisierten Axon involviert ist (Poliak et al. 1999) Nicht nur die Tatsache, dass *CNTNAP2* bereits embryonal exprimiert wird (Abrahams et al. 2007; Alarcón et al. 2008), die Myelinisierung jedoch erst postnatal erfolgt, sondern auch die steigende Zahl der Studien, die einen Zusammenhang des Gens mit ASS berichten, lässt vermuten, dass *CNTNAP2* auch eine entscheidende Rolle in der frühen Phase der Hirnentwicklung spielt (Peñagarikano und Geschwind 2012). Ihm konnte zudem nachgewiesen werden, dass es eine regulatorische Rolle bei der Anhäufung von Kaliumkanälen (Kv1.1/1.2) in der juxtaparanodalen Region der Ranvierschen Schnürringe des Zentralen- und Peripheren Nervensystems spielt (Abbildung 2). Obwohl bereits gezeigt werden konnte, dass die

Kaliumkanäle in dieser Region nicht an der aktiven Erregungsweiterleitung beteiligt sind, wurde dennoch nachgewiesen, dass sie das internodiale Ruhepotential aufrechterhalten (Poliak 2003). Eine Beteiligung bei der Entwicklung und exakten Koordinierung spezifischer Hirnfunktionen kann ihnen somit nicht abgesprochen werden.

Abbildung 2: Schematische Darstellung des strukturellen Aufbaus der juxtaparanodalen Region (a) und seiner Position im myelinisierten Axon (b); nach: Nave 2010.



3.3.2 *CNTNAP2* und Autismus-Spektrum-Störungen

Es hat sich gezeigt, dass bei Verwandten ersten Grades subklinische autistische Verhaltensweisen in sozialen, kognitiven und verbalen Bereichen vorliegen (Bishop et al. 2004) und dass diese subklinischen Verhaltensweisen auch in der allgemeinen Bevölkerung unabhängig von einander vererbt werden (Ronald et al. 2006). Dies setzt jedoch voraus, dass diese Verhaltensweisen durch Veränderungen in bestimmten Hirnregionen hervorgerufen und durch unterschiedliche Faktoren kontrolliert werden, wodurch die autistischen Phänotypen unabhängig von einander in allen Individuen auftreten können (Geschwind 2008). Aufgrund des sehr heterogenen Phänotyps der autistischen Störungen und um die ätiologische Rolle von genetischen Riskovarianten besser zu verstehen werden oft die einzelnen Phänotypen wie z.B. Sprachentwicklung auch unabhängig vom Krankheitsbild untersucht. In Abschnitt 3.2 wurde eine signifikante Rolle von genetischen Varianten von *CNTNAP2* in der Ätiologie von ASS gezeigt. Die möglichen pathologischen Mechanismen der Varianten sind aber weitgehend ungeklärt. Analysen in gesunden Probanden sowie Patienten mit ASS oder spezifischer Sprachentwicklungsstörung (SSES) konnten *CNTNAP2* Varianten sowohl mit Sprachentwicklung als auch mit Neuronaler Konnektivität in Verbindung bringen (Peñagarikano und Geschwind 2012) und sollen nun hier im Detail aufgezeigt werden.

3.3.2.1 *CNTNAP2* und Sprache

Aufgrund des sehr heterogenen Phänotyps der autistischen Störungen wurde der Sprachphänotyp auch unabhängig vom Krankheitsbild untersucht. So konnte gezeigt werden, dass *CNTNAP2* nicht nur bei ASS, sondern auch in vielen anderen Krankheitsbildern den Sprachphänotypen modulierend

beeinflusst. Interessanterweise clustern die meisten der speziell mit dem Sprachphänotypen in ASS oder spezifischen Sprachentwicklungsstörungen (SSES) assoziierten Varianten in einer sehr begrenzten intronischen Region zwischen Exon 13 und 14 (Überblick in: Peñagarikano und Geschwind 2012).

In einer australischen Stichprobe („Raine study“) wurde z.B. die Beteiligung von *CNTNAP2*-Polymorphismen aus dieser Region an spezifischen Sprachentwicklungsstörungen (SSES) in der Normalbevölkerung nachgewiesen. Die Autoren vermuten hierbei, dass einzelne Varianten mit anderen Risikofaktoren interagieren und so die Anfälligkeit für SSES oder ASS erhöhen (Whitehouse et al. 2011). Interessant ist auch, dass bei ASS Patienten gezeigt werden konnte, dass für das Major-C-Allel der *CNTNAP2*-Variante rs2710102 eine Assoziation mit dem Sprachphänotypen „Sprechen erster Worte“ vorliegt (Alarcón et al. 2008). Hingegen wurde bei Kindern mit Dyslexie festgestellt, dass Träger des entgegen gesetzten Minor-T-Allels der gleichen Variante bei einem Non-word Repetition Test, in dem sie Phantasieworte mit 1-5 Silben wiederholen müssen, deutlich schlechter abschneiden, als Major-C-Allel Träger (Peter et al. 2011). Auch bei Patienten mit einer spezifischen Sprachentwicklungsstörung (SSES) konnte bereits für mehrere in unmittelbarer Nähe des SNP rs2710102 liegende Varianten, eine Assoziation mit eben diesem „Non-word Repetition“ Phänotyp nachgewiesen werden (Vernes et al. 2008; Newbury et al. 2011). Auch zeigte sich, dass selbst bei Sprachdefiziten wie Sprachverzögerung (Poot et al. 2010; Sehested et al. 2010) und Stottern (Petrin et al. 2010) komplexe chromosomale Änderungen (z.B. Deletionen) in dem Gen vorliegen können. Die Vermutung, dass *CNTNAP2* an der Sprachentwicklung beteiligt ist, wird dadurch bestärkt, dass bei der Blockierung des Transkriptionsfaktors „Forkhead box P2“ (FOXP2), der im Intron 1 von *CNTNAP2* binden kann und so die Expression von *CNTNAP2* direkt reguliert (Vernes et al. 2008), ebenfalls Sprachstörungen entstehen (Lai et al. 2001; MacDermot et al. 2005).

Auch wenn es im Tierreich bisher keine direkten parallelen zur menschlichen Sprache gibt, zeigen dennoch einige Spezies eine erlernte vokale Kommunikation bei der eine akustisch geleitete, vokale Imitation zu beobachten ist, was wiederum als Voraussetzung für die eigentliche Sprachentwicklung angesehen wird (Peñagarikano und Geschwind 2012). Eine dieser Spezies mit einer gut charakterisierten, erlernten vokalen Kommunikation (Williams 2004) sind die bei den Singvögeln einzugliedernden Zebrafinken (*Taeniopygia guttata*). Bei ihnen konnte bereits in Hirnregionen, die an der Ausbildung des Gesangs beteiligt sind, eine erhöhte Expression von *CNTNAP2* nachgewiesen werden. Diese Regulation des Gens zeigte sich interessanterweise nur bei den männlichen, nicht

aber bei den weiblichen Tieren, was hinsichtlich des Gesangs mit dem Geschlechtsdimorphismus dieser Spezies (Nottebohm und Arnold 1976) übereinstimmt. Zudem zeigen sich insbesondere in der striatalen Region, die bei den Finken als Area X bezeichnet wird und an der Ausbildung des Gesangs beteiligt ist, Parallelen in der Expression des *CNTNAP2* regulierenden Transkriptionsfaktors FOXP2 zwischen den Gehirnen von Zebrafinken und dem sich entwickelnden menschlichen Gehirn (Teramitsu et al. 2004). In Abhängigkeit des aktuellen Lernprozesses konnte gezeigt werden, dass die Regulation der FOXP2 Expression in den entsprechenden Hirnregionen dynamisch reguliert wird. So wird das FOXP2 Gen in jungen, noch im Lernprozess befindlichen Zebrafinken hochexprimiert und nimmt bei älteren Artgenossen, die bereits aus dem Lernprozess hinaus sind und deren Gesang in einer gleich bleibenden, stereotypen Form abläuft, nachweisbar ab (Teramitsu 2006; Teramitsu et al. 2010). Erfolgt während der Lernphase der Singvögel jedoch ein vollständiger Knockdown von FOXP2 in AreaX, führt dies zu unvollständigen und/oder einer fehlerhaften Reproduktion des ursprünglichen Liedgutes (Haesler et al. 2007). Es wird vermutet, dass die Aufgabe des TF FOXP2 bei den Zebrafinken in der Kontrolle der vokalen Ausbildung und somit in der Regulation der Expression seiner Zielgene wie *CNTNAP2* liegt (Überblick in: Peñagarikano und Geschwind 2012).

3.3.2.2 *CNTNAP2* und neuronale Konnektivität

Einer der wenigen Wege um in-vivo die funktionelle Verknüpfung zwischen Genotyp und Phänotyp am menschlichen Gehirn zu erforschen ist die Durchführung einer nicht invasiven Neuro-Imaging-Studie.

Auf diese Weise konnte unter Anwendung der Magnetresonanztomographie (MRT) bereits gezeigt werden, dass durch genetische Varianten Veränderungen in der Hirnstruktur und -funktion entstehen können, die oftmals mit psychiatrischen Krankheitsbildern wie ASS einhergehen. In einer in der Normalbevölkerung durchgeführten Studie wurde so die Morphologie der weißen und grauen Substanz bei Trägern des T-Allels der *CNTNAP2*-Variante rs7794745 untersucht (Tan et al. 2010). Es zeigte sich, dass die homozygoten Risikoallelträger (rs7794745T) in Hirnbereichen, die bereits mit ASS in Verbindung wurden wie dem Cerebellum, dem Gyrus fusiformis und dem occipitalen und frontalen Cortex eine signifikante Reduzierung des Volumens der weißen und grauen Substanz aufweisen, obwohl bei den getesteten Personen keine erkennbaren Verhaltensstörungen vorlagen. Interessanterweise wurde für diese Variante bereits in einer verhergehenden Assoziationsstudie gezeigt, dass sie das Risiko für ASS erhöht (Arking et al. 2008).

Für eine weitere *CNTNAP2*-Variante (rs2710102), welche nachweislich den Sprachphänotypen „Sprechen erster Worte“ in ASS negativ beeinflusst (Alarcón et al. 2008) konnte in einer weiteren Neuroimaging-Studie (Scott-Van Zeeland et al. 2010) gezeigt werden, dass das Major-C-Allel im Zusammenhang mit einer abnormalen funktionellen Konnektivität im Frontallappen steht. Bei Lernaufgaben zeigte sich eine Aktivierung des medialen präfrontalen Cortex (mPFC), obwohl diese normalerweise nur die fronto-striatalen-Bereiche ansprechen. Zudem zeigte sich bei einer Analyse der funktionellen Konnektivität zwischen dem mPFC und den posterior corticalen Regionen eine erhöhte lokale Konnektivität bei den Personen, die das C-Risikoallel tragen. Im Vergleich dazu konnte bei den Nicht-Risikoallelträgern (T) überwiegend Verknüpfungen der anterior und posterior gelegenen Hirnregionen nachgewiesen werden (Scott-Van Zeeland et al. 2010). Dies könnte ein Hinweis darauf sein, dass im Gehirn der Risikoallelträger vermehrt unterentwickelte Verknüpfungen vorliegen, da der Rückgang lokaler Verbindungen verbunden mit der Erhöhung von Verknüpfungen mit weiter entfernt liegenden Hirnregionen ein natürlicher Vorgang beim Reifungsprozess des Gehirns ist (Dosenbach et al. 2010). Aufgrund dieser Ergebnisse lässt sich vermuten, dass ASS das Resultat von Fehlentwicklungen während des Reifeprozesses des Gehirns ist, die auf fehlerhaft abgelaufene molekulare Mechanismen zurückzuführen sind (Geschwind und Levitt 2007). So könnten genetische Varianten von *CNTNAP2* zu Fehlverknüpfungen von Hirnregionen führen, die an höheren kognitiven Funktionen beteiligt sind (Pardo und Eberhart 2007; Rubenstein 2011). Diese Fehlverknüpfungen treten in Abhängigkeit vom genetischen Hintergrund auch unabhängig von einer ASS Diagnose auf (Scott-Van Zeeland et al. 2010). *CNTNAP2* Varianten, welche die Konnektivität ändern, können für sich alleine nicht dieses komplexe Krankheitsbild der autistischen Störungen hervorbringen, es wird aber vermutet, dass diese modulierend auf ASS wirken und somit als Risikofaktoren angesehen werden können. (Überblick in: Peñagarikano und Geschwind 2012).

3.3.3 Das Cntnap2-Mausmodell und Autismus-Spektrum-Störungen

Um die Funktion einzelner Gene im Kontext von ASS in-vivo zu erforschen wird gerne auf das Mausmodell (*Mus musculus*) zurückgegriffen, da sie als Säugetiere nah genug mit dem Menschen verwandt sind, um an ihnen genetisch bedingte Verhaltensänderungen in der sozialen Interaktion, der Kommunikation und dem Verhalten zu erforschen. So konnte auch durch die Erstellung von *Cntnap2*-Knockout (KO) Mäusen gezeigt werden, dass die Tiere in den drei autistischen Kernbereichen (Kommunikation, soziale Interaktion und Verhalten) Defizite aufweisen. Zusätzlich waren bei ihnen Symptome wie Hyperaktivität und epileptische Anfälle zu beobachten, die auch

bereits bei Patienten mit FCD nachgewiesen werden konnten (Peñagarikano et al. 2011). Auch neuropathologisch zeigten die KO Mäuse Ähnlichkeiten zu den FCD Patienten. So konnten bei ihnen Fehler bei der Migration der corticalen Projektionsneuronen festgestellt werden, was die bereits bei FCD Patienten beobachteten Abweichungen bei der neuronalen Migration bestätigt (Strauss et al. 2006). Der außerordentliche Grad der Organisation des Säugetiergehirns zeigt zudem, in welcher präzisen Art und Weise die Migrationsbewegungen während der Entwicklung des Gehirns eigentlich ablaufen sollten. Auch können Störungen der neuronalen Migration während der Hirnentwicklung unter anderem zu Epilepsien, geistigen Behinderungen und anderen neurologischen Entwicklungsstörungen führen (Guerrini und Barba 2010). Abgesehen von den Veränderungen während der neuronalen Migration zeigten die *Cntnap2*-KO Mäuse eine Reduzierung der GABAergen Interneuronen (Peñagarikano et al. 2011). Das eröffnete eine neue Perspektive auf die Funktion des Gens *CNTNAP2*, da die Untersuchung seiner funktionellen Rolle bisher auf die exzitatorischen Pyramidalzellen beschränkt war. Es ist durchaus möglich, dass der das Verhalten betreffende Phänotyp die Folge eines Ungleichgewichts in der exzitatorischen und inhibitorischen Erregungsweiterleitung ist, dessen Beteiligung an der Pathophysiologie von ASS bereits diskutiert wurde (Rubenstein und Merzenich 2003). Die in den KO Mäusen vorliegende abnormale Hirnstruktur wird zusätzlich durch die höchst unsynchron ablaufende corticale Aktivität verdeutlicht (Peñagarikano et al. 2011), was wiederum als pathophysiologischer Mechanismus in ASS gilt (Uhlhaas und Singer 2006). Besonders interessant ist, dass auch die Dysfunktion der GABAergen Interneuronen, welche im Mausmodell ja reduziert vorliegen, als möglicher Auslöser für Abweichungen in der neuronalen Synchronie diskutiert werden (Bonifazi et al. 2009).

Insgesamt zeigen die Daten aus den Mausmodellen, dass die Desorganisation der neuronalen Kreisläufe Veränderungen in der Informationsverarbeitung hervorruft. Die dem beobachteten Mausphänotyp zu Grunde liegende Neuropathologie ist jedoch weiterhin nicht bekannt. Es sind daher weitere Untersuchungen der Struktur von neuronalen Netzwerken und der Funktion der Interneuronen in *Cntnap2*-KO Mäusen nötig, da nur so die Aufklärung der krankheitsverursachenden Mechanismen von ASS voran getrieben werden können (Überblick in: Peñagarikano und Geschwind 2012). In Summe kann auch durch diese Studien eindrucksvoll gezeigt, dass *CNTNAP2* eine entscheidende Rolle bei der Entstehung von ASS spielt.

3.3.4 Genexpressionsstudien und *CNTNAP2*

Um herauszufinden, in welchen Hirnbereichen *CNTNAP2* während der Entwicklung des Gehirns exprimiert wird, haben Alarcón und Kollegen (2008) eine Expressionsstudie an zwei weiblichen und fünf männlichen Gehirnen des Menschen durchgeführt. Hierbei konnten sie zeigen, dass *CNTNAP2* vor allem in jenen Hirnregionen angereichert vorliegt, die auch an der Sprachentwicklung beteiligt sind. In einer anderen Expressionsstudie wurde zudem gezeigt, dass in Föten der 18-22 Gestationswoche die *CNTNAP2*-Expression in Abhängigkeit des Transkriptionsfaktors FOXP2 deutlich herunterreguliert wird (Vernes et al. 2008). Diese Runterregulierung von *CNTNAP2* während der neuronalen Differenzierung konnte in humanen neuronalen Vorläuferzellen einer anderen Studie bestätigt werden (Konopka et al. 2012). Basierend auf diesen Ergebnissen hat Frau Denise Haslinger (M.Sc.) ebenso gezeigt, dass auch in einem SH-SY5Y Zellmodell für neuronale Differenzierung die Genexpression von *CNTNAP2* während der Differenzierungsphase (72 Stunden nach Induktion der Zelldifferenzierung) zunächst deutlich herunterreguliert wird und zum Ende der Differenzierung wieder ansteigt (Daten sind bisher nicht publiziert). Auch Kang und Kollegen (2011) konnten in ihrer Studie zeigen, dass die *CNTNAP2*-Expression in gesunden Individuen prä- und perinatal erfolgt, aber im frühen Kindesalter (1-6 Jahren) stark runterreguliert wird. Zudem konnten sie nachweisen, dass nach dieser Runterregulierung eine Reexpression von *CNTNAP2* in fast allen Hirnregionen erfolgt. Interessanterweise sind das Striatum und der Cerebellar Cortex, zwei Hirnregionen, deren Beteiligung an ASS (Haas et al. 1996; Langen et al. 2009) und am hierarchischen Lernen während des Spracherwerbs bereits diskutiert wurde (Strick et al. 2009; Desrochers und Brade 2012), von der lebenslangen Reexpression ausgenommen.

3.3.5 Promotorvarianten in der *CNTNAP2*-Promotorregion

Die 5'-liegende Promotorregion gilt als primäres Element der Genregulation. Es ist daher nicht verwunderlich, dass Promotorvarianten die dort lokalisierten Transkriptionsfaktorbindestellen modifizieren können und so einen Einfluss auf die Genexpression haben (Martin et al. 2012). Trotz der offensichtlichen funktionellen Relevanz von genetischen Varianten in der Promotorregion wurden die Varianten in der *CNTNAP2*-Promotorregion bisher weder im Zusammenhang mit ASS charakterisiert, noch wurde ihr möglicher funktioneller Einfluss auf die Transkription genauer untersucht. Eine detaillierte Untersuchung der Varianten der *CNTNAP2*-Promotorregion auf genetischer und funktioneller Ebene könnte zeigen, ob diese eine entscheidende Rolle in der Ätiologie von ASS haben und deren möglichen Pathomechanismus aufklären.

3.4 Ziel der Studie

Basierend auf den unterschiedlichsten Studienansätzen konnte in den letzten Jahren immer wieder gezeigt werden, dass Varianten des *CNTNAP2* Gens an der Entstehung von ASS beteiligt sind und auch den Sprachphänotyp bei ASS modulierend beeinflussen. Trotz der allgemein bekannten funktionellen Relevanz von Promotorvarianten wurde der *CNTNAP2*-Promotorregion bisher keine Beachtung geschenkt. In dieser Arbeit wurde aufgrund der in der Einleitung dargelegten Studien vermutet, dass Promotorvarianten über eine veränderte Transkriptionseffizienz des *CNTNAP2*-Gens modulierend auf ASS und den Sprachphänotyp wirken.

Ziel dieser Studie war es daher, bekannte und unbekannte Varianten in der *CNTNAP2*-Promotorregion zu detektieren und deren Einfluss auf die Sprachentwicklung sowie deren funktionellen Einfluss auf die Aktivität des Promotors zu untersuchen. Die Ergebnisse dieser Studie erlauben einen Rückschluss auf die funktionellen Mechanismen der mit ASS assoziierten Varianten der *CNTNAP2*-Promotorregion.

Im Rahmen dieser Arbeit wurden in einer 236 Familien umfassenden Detektionsstichprobe, die phänotypisch gut charakterisierten ist, bekannte und unbekannte Varianten mittels direkter Sequenzierung nachgewiesen. Basierend auf den Ergebnissen der Sequenzierung wurden anschließend die häufigsten und/oder mit ASS assoziierten Varianten in einer Replikationsstichprobe von 356 Familien genotypisiert. Durch eine weitere Assoziationsanalyse sollte überprüft werden, ob die Ergebnisse der Detektionsstichprobe replizierbar sind, bzw. ob aufgrund der Vergrößerung der Stichprobe weitere Assoziationen der jeweiligen Varianten mit ASS nachgewiesen werden können. Eine Assoziation mit dem Sprachphänotypen wurde durch eine Regressionsanalyse getestet. Der funktionelle Einfluss der einzelnen Varianten wurde in-silico und in-vitro charakterisiert. In-silico wurde getestet, ob die jeweiligen Promotorvarianten innerhalb von Bindesequenzen hirnspezifischer Transkriptionsfaktoren liegen. In der abschließenden in-vitro Analyse wurde mittels Luciferase-Assay der Einfluss der Varianten auf die Aktivität des *CNTNAP2*-Promotors untersucht. Hierbei wurde die Transkriptionseffizienz speziell in fibroblastoiden mitotisch aktiven Zellen (HEK293T) sowie in neuroblastoiden Zellen (SH-SY5Y) vor und während der Differenzierung untersucht, um eventuelle neuronal spezifische Effekte zu detektieren.

4 Material

4.1 Geräte

Gelelektrophorese

<i>Horizontal</i>	Compact S/XS	Biometra GmbH
	Compact M	Biometra GmbH
	Compact L/XL	Biometra GmbH
	EasyPhore Medi	Biozym Scientific GmbH
<i>Vertikal</i>	Multigel-Long	Biometra GmbH
Powersupply	Power Pack P25 T	Biometra GmbH
	Power Source	VWR
Taumel-Wipptisch	WT17	Biometra GmbH
UV-Dokumentation	Alphamager HP	Cell Bioscience, USA

Pipetten

Handdispenser	Distriman®	Gilson
	Multipette® plus	Eppendorf
Mehrkanal	Research® plus	Eppendorf
Pipettierhilfe	Accurpette	VWR
Variabel	Research® plus	Eppendorf

Zellkultur

Inkubator	HeraCell 150i CO ₂	Thermo Scientific
Luminometer	TECAN Infinite M200	Tecan Deutschland GmbH
	GloMax®-Multi Detection System	Promega GmbH
Trinokularmikroskop	AE31	Motic®
	Kamera Moticam 3 (CMOS)	Motic®
Wasserbad	für 4-12l	VWR
Werkbank	MSC Advantage 1.2	Thermo Scientific
Zählkammer	Neubauer improved	VWR

Zentrifugen

Universalzentrifuge	Heraeus Megafuge 16R	Thermo Scientific
Mikrozentrifuge	Heraeus Fresco 17/21	Thermo Scientific
	Heraeus Pico 17/21	Thermo Scientific
Minizentrifuge	Sprout	Biozym Scientific GmbH

4.2 Sonstige Geräte und Verbrauchsmaterialien

Einmalpipetten	Serologische Pipetten 2ml / 5ml	Sarstedt
	Serologische Pipetten 10ml / 25ml	Costar
Filterspitzen	Biosphere® 10µl	Sarstedt
	Biosphere® 100µl	Sarstedt
	Safe Seal Tips Professional 1000µl	Biozym Scientific GmbH
	Safe Seal Tips Premium 5ml	Biozym Scientific GmbH
Kapillarspitzen	MidiTips Premium 200µl	Biozym Scientific GmbH
Kulturflaschen	CELLSTAR® Filter Top 25cm ²	Greiner
	CELLSTAR® Filter Top 75cm ²	Greiner
PCR-Maschine	LabCycler	Sensoquest
Photometer	NanoPhotometer™	Implen
Reaktionsgefäße	Multiply®-µStrip Pro 8er Kette	Sarstedt
	Mikro-Schraubröhre 1,5ml, PP	Sarstedt
	SafeSeal Gefäß 1,5ml	Sarstedt
	Mikro-Schraubröhre 2ml, PP	Sarstedt
	96well Multiply®-PCR Platte	Sarstedt
	Zentrifugenröhre 15ml	Greiner
	Zentrifugenröhre 50ml	Sarstedt
	Röhre 13ml; mit Belüftungsstopfen	Sarstedt
	24-well, F-Boden	Sarstedt
	CytoOne TC 12-well, F-Boden	Starlab
Zellkulturplatten		

4.3 Chemikalien

Agarose High Resolution	Carl Roth GmbH
Aqua ad iniectabilia	B. Braun Melsungen AG
Aqua _{bidest}	Carl Roth GmbH
Ammoniumperoxidsulfat (APS)	Carl Roth GmbH
Ammoniumsulfat (NH ₄) ₂ SO ₄	Applichem
Ampicillin Natriumsalz	Carl Roth GmbH
B-27 supplement (50x)	Life Technologies GmbH
Brain-derived neurotrophic factor human (hBDNF)	Sigma-Aldrich
Dibutyl-cAMP C ₁₈ H ₂₃ N ₅ O ₈ PNa	Sigma-Aldrich
Desoxynucleosidtriphosphate (2mM each)	Fermentas GmbH
DH5α TM Competent Cells (#18265-017)	Life Technologies GmbH
Dikaliumhydrogenphosphat K ₂ HPO ₄	Applichem
Dimethylsulfoxid (DMSO)	Carl Roth GmbH
DMEM+GlutaMAX	Life Technologies GmbH
Dulbecco's PBS (1x)	PAA Laboratories GmbH
EDTA-Dinatriumsalz C ₁₀ H ₁₄ N ₂ Na ₂ O ₈ x 2H ₂ O	Applichem, Darmstadt
Ethanol absolut	Sigma-Aldrich
Fetal Bovine Serum (FBS)	Life Technologies GmbH
GelRed TM (10.000x in H ₂ O)	Biotium, Hayward, USA
GlutaMax (100x)	Life Technologies GmbH
Glycerin, ROTIPURAN® (≥99,5 %)	Carl Roth GmbH
Isopropanol	Sigma-Aldrich
Kaliumchlorid KCl (≥99.0%)	Sigma-Aldrich
Kaliumdihydrogenphosphat KH ₂ PO ₄	Carl Roth GmbH
Kanamycin Sulfat	Sigma-Aldrich
LB-Agar (Lennox)	Carl Roth GmbH
LB-Medium (Lennox)	Carl Roth GmbH
Magnesiumsulfat MgSO ₄	Applichem
Metafectene® Pro	Biontex Laboratories GmbH
Natriumacetat C ₂ H ₃ NaO ₂ (3M)	Applichem
Natriumchlorid NaCl	Applichem

Natriumlaurylsulfat $C_{12}H_{25}NaO_4S$ (SDS)	Applichem
Neurobasal-A-Medium (ohne Phenolrot)	Life Technologies GmbH
Opti-MEM Reduced Serum Medium	Life Technologies GmbH
Orange G	Applichem
Penicillin/Streptomycin (100x)	PAA Laboratories GmbH
Phenol-Chloroform; 50:50; pH 8,0	Carl Roth GmbH
Proteinase K (20mg/ml)	Applichem
Retinsäure $C_{20}H_{28}O_2$	Sigma-Aldrich
Rotiphorese® Gel 40 (19:1)	Carl Roth GmbH
SeaKem® LE Agarose	Lonza GmbH
Sodium Pyruvat (100mM)	Life Technologies GmbH
Taurin	Carl Roth GmbH
TEMED	Carl Roth GmbH
TE-Puffer (pH 8,0)	Applichem
Trinatriumcitrat Na_3 -Citrat	Merck KGaA
Tris-Base $C_4H_{11}NO_3$	Carl Roth GmbH
Trypanblau-Lösung 0,4%	Sigma-Aldrich
Trypsin EDTA 0,05%	Life Technologies GmbH
X-Gal in DMF (40mg/ml)	Biological, Massachusetts, USA
X-tremeGENE HP DNA Transfection Reagent	Roche Diagnostics GmbH
Xylencyanol	Merck KGaA

4.4 Puffer, Gele und Lösungen

4.4.1 DNA-Isolierung aus Wangenabstrichen

Lyse-Puffer

10mM Tris-Base

10mM EDTA x $2H_2O$

100mM NaCl

2%(w) SDS

Aqua_{bidest}

pH 8,0 +/- 0,5; autoklavieren

4.4.2 Polymerasekettenreaktion

Oligonukleotide

Alle Primer wurden von der Firma Sigma-Aldrich (Steinheim) bezogen und als 100µM Stammlösung in TE-Puffer angesetzt und gelagert (20°C). Aus der Stammlösung wurde eine 10µM Arbeitslösung in Aqua ad iniectabilia angesetzt. Detaillierte Informationen zu den PCR-Bedingungen sind bei den jeweiligen Methoden aufgelistet.

4.4.3 Gelelektrophorese

TTE-Puffer (20x)	Ladepuffer (5x)	Polyacrylamid Gel (7,5 %)
1,8M Tris-Base	30%(v) Glycerin	75%(v) Aqua _{bidest}
0,6M Taurin	50mM EDTA x 2H ₂ O	5%(v) 20x TTE
0,02M EDTA x 2H ₂ O	0,001%(w) Xylencyanol	18,75%(v) Rotiphorese® Gel 40 (19:1)
Aqua _{bidest}	0,001%(w) Orange G	1%(v) APS (10%(w))
	Aqua _{bidest}	0,1%(v) TEMED

4.4.4 Klonierung

4.4.4.1 Antibiotika

Ampicillin Natriumsalz

Aus dem Ampicillin Natriumsalz wurde eine 100mg/ml Stammlösung angesetzt, welche nach der Sterilfiltration bei -20°C gelagert wurde. Das Antibiotikum wurde den jeweiligen Nährmedien mit einer Endkonzentration von 100µg/ml zugegeben.

Kanamycin Sulfat

Aus dem Kanamycin Sulfat wurde eine 50mg/ml Stammlösung angesetzt, welche nach der Sterilfiltration bei -20°C gelagert wurde. Das Antibiotikum wurde den jeweiligen Nährmedien mit einer Endkonzentration von 50µg/ml zugegeben.

4.4.4.2 Dauerkultur: Hogness Modified Freezing Medium (10x; HMFM)

Das 10x HMFM kann direkt zum LB-Medium (Lennox) gegeben werden. Dies ermöglicht das direkte Einfrieren der Dauerkultur ohne erneute Zugabe von Glycerin.

Teil 1

5mM MgSO₄
 20mM Na₃-Citrat
 70mM (NH₄)₂SO₄
 45%(v) Glycerin
 Aqua_{bidest}

Teil 2

1,3M K₂HPO₄
 0,66M KH₂PO₄
 Aqua_{bidest}

Teil 1 und 2 werden getrennt von einander angesetzt und autoklaviert. Anschließend werden sie im Verhältnis 5:1 (z.B.: Teil 1: 100ml, Teil 2: 25ml) gemischt.

4.4.5 Zellkultur***Standardmedium***

DMEM+GluMAX
 10%(v) Fetal Bovine Serum (FBS)
 1x Sodium Pyruvate
 1x Penicillin/Streptomycin (100x)

Einfriermedium

60%(v) DMEM+GluMAX
 30%(v) FBS
 1x Sodium Pyruvate
 10%(v) DMSO

Neurobasal-A-Medium mit hBDNF/RA

Neurobasal-A-Medium (ohne Phenolrot)
 1x Glutamax (100x)
 1x B-27 supplement (50x)
 2mM Dibutyryl-cAMP
 50ng/ml hBDNF
 20mM Kaliumchlorid
 1x Penicillin/Streptomycin (100x; optional)
 10µM Retinsäure (0,1M)]

Stocklösung Retinsäure

0,1M Retinsäure (300,44g/mol)
 In DMSO

 → Retinsäure ist wasserunlöslich
 Lösung entspricht 30mg/ml
 Stocks lichtgeschützt bei -80° aufbewahren

4.5 Probenmaterial

Untersucht wurden insgesamt 2170 Personen, bestehend aus 756 Probanden mit der Diagnose einer Autismus-Spektrum-Störung (ASS) nach ADI-R und/oder ADOS-G sowie und deren Familienangehörigen ohne Diagnose. Alle teilnehmenden Probanden stellten entweder Blut oder Wangen-

abstriche zur Gewinnung genomischer DNA zu Verfügung. Alle Patienten und/oder deren gesetzliche Vertreter wurden über den Ablauf der Studie informiert und haben diesem zugestimmt. Die Ethikvoten der verschiedenen Zentren für die Blutentnahme und die anschließende genetische Untersuchung liegen unter den folgenden Geschäftsnummern vor: 162/99 (Frankfurt; Poustka); 73/04 (Homburg; Freitag); 237/09 (Frankfurt; Freitag).

Die Probanden wurden an der Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie des Universitätsklinikums des Saarlandes sowie der Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes und Jugendalters (KJPPP) der Johann Wolfgang Goethe-Universität rekrutiert. Die Diagnosen wurden unter Anwendung des „Diagnostischen Interviews für Autismus - Revidiert“ (ADI-R; Lord et al. 1994) und der „Diagnostischen Beobachtungsskala für Autistische Störungen“ (ADOS; Lord et al. 2000) gestellt. Ebenso erfolgte eine neurologische Untersuchung der Patienten sowie eine Bestimmung des IQ-Wertes unter Anwendung des „Hamburg-Wechsler-Intelligenztest für Kinder“ (WISC/WISE; HAWIK-R/HAWIE-R; Wechsler 1949), oder des „Columbia Mental Maturity Scale“ (CMMS; Burgemeister et al. 1954). Patienten mit einem IQ-Wert < 35, einem Geburtsgewicht < 1000g oder einer bekannten chronischen Erkrankung (mit Ausnahme von Epilepsie) wurden nicht in die Studie aufgenommen. Weitere Ausschlusskriterien waren Blindheit, Taubheit, karyotypische Anomalien, Tuberöse Sklerose oder das Vorliegen einer zerebraler Kinderlähmung oder eines Fragilen-X Syndroms.

Von den insgesamt 756 Patienten erfüllten 592 alle Kriterien, bestehend aus 510 männlichen und 82 weiblichen Patienten aus insgesamt 592 nicht verwandten Familien. Diese unterteilen sich in 492 Eltern-Kind-Trios, 73 Eltern-Kind-Duos sowie 27 Einzelkinder. Zusätzlich waren DNA Proben von 96 Geschwisterkindern sowie 23 sonstigen Angehörigen verfügbar. Insgesamt bestand die finale Stichprobe aus 1769 Subjekten.

Bei 518 der 592 Patienten lag die Diagnose eines frühkindlichen Autismus vor, von denen 317 einen $IQ \geq 70$ (HFA = High-Functioning-Autismus) bzw. 201 einen $IQ \leq 70$ (LFA = Low-Functioning-Autismus) aufwiesen. Bei den restlichen 74 Patienten wurde eine Spektrum-Störung diagnostiziert, von denen wiederum 56 Patienten einen Asperger und 18 einen atypischen Autismus (PDD-NOS) aufwiesen.

Das Gesamtset wurde in eine Detektionsstichprobe (236 Familien) und eine Replikationsstichprobe (356 Familien) unterteilt. Die Detektionsstichprobe bestand aus allen Proben, die dem Molekulargenetischen Labor der KJPPP (Frankfurt) bis November 2010 zur Verfügung standen. Die Replikationsstichprobe bestand aus Proben, die vor der Gründung (04.2009) des Molekulargenetischen Labors der KJPPP unter der Leitung von Prof. Dr. Poustka in Frankfurt rekrutiert und am Deutschen Krebsforschungszentrum (DKFZ) in Heidelberg aufgereinigt bzw. bis Januar 2011 dort gelagert wurden. Alle weiteren Proben, die im Zeitraum von Dezember 2010 bis Dezember 2011 in den Bestand des Molekulargenetischen Labors der KJPPP (Frankfurt) eingegangen waren, wurden in die Replikationsstichprobe eingeschlossen.

5 Methoden

5.1 Studiendesign

Nach vorangegangener DNA-Extraktion aus Blutproben oder Wangenabstrichen sollten bei einem Screening in einer Detektionsstichprobe durch direkte Sequenzierung neue und bereits beschriebene Varianten innerhalb der *CNTNAP2*-Promotorregion nachgewiesen werden. Die detektierten Varianten wurden auf Assoziation zu ASS und ASS-spezifischen Phänotypen getestet. In der daran anschließenden Replikationsstudie wurden nur die Varianten genotypisiert, die in der Detektionsstichprobe am häufigsten vorkamen (Genotypfrequenz > 1%) bzw. im Transmission-Disequilibrium-Test (TDT) eine Assoziation mit ASS aufwiesen (p-Wert < 0,05). Die funktionelle Untersuchung der ausgewählten Varianten in der *CNTNAP2*-Promotorregion erfolgte mittels Luciferaseassay. Der Aufbau der Gesamtstudie geht aus Tabelle 1 hervor.

Tabelle 1: Schematischer Aufbau der Gesamtstudie

DNA-Extraktion		
Detektionsstudie → Nachweis neuer und bereits beschriebener Varianten - Sequenzierung - Auswertung	Replikationsstudie → Vergrößerung der Stichprobe und Replikation der bisherigen Ergebnisse - Genotypisierung RFLP (Frankfurt) RT-PCR (Heidelberg) - Auswertung	Funktionelle Analyse → Untersuchung möglicher Effekte einzelner Varianten auf die Promotorfunktion - In-silico Analyse - Luciferaseassay - Auswertung

5.2 Isolierung der DNA aus dem Probenmaterial

Die Isolierung der DNA erfolgte aus Wangenabstrichen und Blut. Nach jeder Extraktion wurde die Konzentration der gewonnenen Stocklösungen bestimmt.

5.2.1 Isolierung der DNA aus EDTA-Blut

Die DNA-Extraktion aus den eingegangenen Blutproben erfolgte mit dem ‚MasterPure™ DNA Purification Kit for Blood Version II‘ (#MB711400) von Epicentre Biotechnologies (Madison, USA) nach dem Standardprotokoll, welches wie folgt angepasst wurde:

Bei Erhalt von 2x9ml EDTA-Blut einer Person wurde die DNA aus beiden Monovetten zunächst getrennt aufgereinigt. Dafür sind die gesamten 9ml Blut eines Röhrchens jeweils in ein 50ml Falcon überführt worden. Nach der Lyse der Erythrozyten mit *Red Cell Lysis Solution* wurden die Leukozyten für 10 Minuten bei 2000g abzentrifugiert. Der Überstand wurde bis auf ca. 500µl abgegossen und das Leukozytenpellet durch vortexen in der Restflüssigkeit gelöst. Die Lyse der weißen Blutzellen erfolgte in *Tissue and Cell Lysis Solution*. Durch Zugabe von RNase A wurde noch vorliegende RNA zerstört. Nach 30 Minuten Inkubation bei 37°C wurden die Proben für 3-5 Minuten auf Eis runtergekühlt, Proteine durch die Zugabe von *MPC Protein Precipitation Reagenz* gefällt und durch Zentrifugation für 20 Minuten bei 4696g präzipitiert. Der Überstand wurde zu 22,5ml vorgelegtem Isopropanol gegeben, das Tube 30-40mal invertiert und die DNA durch Zentrifugation für 20 Minuten bei 4°C und 4696g gefällt. Nach dem Verwerfen des Überstandes und nach 2maligem Waschen der Pellets mit je 2ml 70%igem Ethanol (2 Minuten; 4696g) wurden die DNA-Pellets aus den bisher getrennt bearbeiteten Proben in einem frischen 2ml Reaktionsgefäß vereint, für 10 Minuten mit geöffnetem Deckel getrocknet und abschließend in 1ml TE-Puffer gelöst. Damit sich die DNA besser löst, wurde das Probenröhrchen zunächst für eine Stunde bei 65°C inkubiert, kurz gevortext und vor der Qualitäts- und Konzentrationsbestimmung für mindestens 2 Tage im Kühlschrank gelagert. Nach der Bestimmung von Qualität und Konzentration der DNA mit dem Implen ‚Nanophotometer‘ wurde eine Arbeitslösung mit einer Konzentration von 20ng/µl angesetzt, welche bei -20°C aufbewahrt wird. Die Stocklösung lagert dauerhaft bei -80°C. In sämtliche PCR-Ansätze wurde die DNA mit einer Konzentration von 20ng/µl eingesetzt.

5.2.2 Isolierung der DNA aus Wangenabstrichen

Die DNA-Extraktion aus Wangenabstrichen erfolgte mittels Phenol-Chloroform-Methode. Alle Zentrifugationsschritte wurden bei 15.000g durchgeführt.

Nach dem Wangenabstrich wurde der Swab in ein 2ml Reaktionsgefäß abgeworfen, mit 750µl Lyse-Puffer und 30µl Proteinase K (20mg/ml) versetzt und für 30 Minuten bei 56°C inkubiert. Unter einen Sicherheitsabzug wurden 750µl Phenol-Chloroform (50:50) dazu pipettiert, das Reagiergefäß gevortext und für 5 Minuten zentrifugiert. Die obere, wässrige Phase, welche die DNA enthält, wurde vorsichtig abgenommen und in ein neues Gefäß überführt. Um die DNA zu fällen wurden 80µl Natriumacetat (3M) und 1 Volumen (750µl) Isopropanol dazugegeben. Das Gemisch wurde gevortext, 15 Minuten bei Raumtemperatur inkubiert und anschließend erfolgte die Pelletierung der DNA durch 15minütige Zentrifugation. Der Überstand wurde verworfen und das Pellet mit 1ml 75% Ethanol gewaschen. Nach 5minütiger Zentrifugation des Reaktionsgefäßes wurde der Ethanol abgezogen, das Pellet für mindestens 10 Minuten bei 56°C mit offenem Deckel auf dem Heizblock getrocknet und in 25µl TE gelöst. Abschließend wurden Qualität und Konzentration der DNA mit dem Implen ‚Nanophotometer‘ bestimmt und eine Arbeitslösung mit einer Konzentration von 20ng/µl angesetzt, welche bei -20°C aufbewahrt wird. Die Stocklösung lagert dauerhaft bei -80°C. In sämtliche PC-Reaktionen wurde die DNA mit einer Konzentration von 20ng/µl eingesetzt.

5.2.3 Konzentrationsbestimmung und Qualitätskontrolle der extrahierten DNA

Um Arbeitslösungen einer bestimmten Konzentration ansetzen zu können, wurde die Gesamtkonzentration der DNA-Probe mit dem Implen ‚Nanophotometer‘ bestimmt.

Dafür wurden 1-3µl der in TE-Puffer oder Aqua ad iniectabilia gelösten DNA auf die Meßvorrichtung aufgetragen und anschließend die Extinktion der unbekannten Probe bei verschiedenen Wellenlängen gemessen. Die Konzentration von doppelsträngiger DNA errechnet sich aus der Extinktion E260 nach folgender Formel:

$$C = E_{260} \times f \times c$$

C = Konzentration der gemessenen Probe in µg/ml

E260 = Extinktion bei 260nm

f = Verdünnungsfaktor

c = Konzentrationstandard /Extinktionstandard

für doppelsträngige DNA: c = 50 µg/ml

Da das Absorptionsmaximum von Proteinen bei 280nm liegt, kann der Quotient aus E260/E280 als Maß für die Reinheit der DNA genutzt werden. Der Wert von reiner DNA ($E260/E280 > 1,8$) wird durch Proteinverunreinigungen verringert.

Das Absorptionsmaximum von Verunreinigungen wie Peptide, Phenole, Kohlehydrate oder aromatische Verbindungen liegt hingegen bei 230nm. Bei reinen DNA-Proben sollte der Quotient aus E260/E230 größer 2,0 sein.

E230 = Absorptionsmaximum von organischen Substanzen

E260 = Absorptionsmaximum von DNA

E280 = Absorptionsmaximum vieler Proteine

5.3 Analyse der Detektionsstichprobe

Nach der Isolierung der DNA aus dem Probenmaterial wurden sämtliche Proben der Detektionsstichprobe mittels direkter Sequenzierung analysiert, um bereits bekannte Varianten in unserer Stichprobe zu bestätigen und bisher nicht beschriebene Varianten nachzuweisen. Hierzu wurde die *CNTNAP2*-Promotorregion zunächst mittels Polymerasen-Ketten-Reaktion (PCR) amplifiziert. Das PCR-Produkt wurde aufgereinigt in die Sequenzierreaktion eingesetzt und die dabei entstandenen Fragmente mittels Kapillarelektrophorese aufgetrennt. Basierend auf den Ergebnissen der Sequenzanalyse erfolgte die Auswahl der Varianten, die in der Replikationsstichprobe und der funktionellen Studie weiter untersucht werden sollten.

5.3.1 Sequenzierung der *CNTNAP2*-Promotorregion

Das Kettenabbruch-Verfahren nach Sanger (Sanger et al. 1992) ist eine der am häufigsten genutzten Methoden, um die exakte Basenabfolge in einem DNA-Strang zu ermitteln. Bei der Cycle Sequencing Reaktion dient das doppelsträngige PCR-Produkt der Polymerase als Matrize der Polymerase für die Synthese eines komplementären Strangs. Durch den Einbau von Didesoxynukleosidtriphosphaten (ddNTPs), welche neben Desoxynukleosidtriphosphaten (dNTPs) zusätzlich in den Reaktionsansatz gegeben werden, kommt es zu einem Abbruch der Amplifikation. Die entstandenen DNA-Fragmente können anschließend mittels Kapillarelektrophorese größenspezifisch aufgetrennt werden.

5.3.1.1 Amplifikation der Promotorregion für die Sequenzierung

Für die Sequenzierung des *CNTNAP2*-Promotors wurde die Region +586 bis -710bp (NM_014141.5) mit den Primern CasProm-F-34712024 und CasProm-R-2462603 amplifiziert. Die Auswahl der Sequenz erfolgte aufgrund der in ENCODE annotierten Transkriptionsfaktorbindestellen und regulatorischen Elemente in der 5' Region vor den kodierenden Sequenzen des Gens.

Die PCR erfolgte nach Herstellerangabe mit der TopTaq™ Polymerase (#200205; Qiagen) unter Zugabe von Q-Solution. Vor der Sequenzierung wurden die Amplikons mit dem GeneJET™ PCR Purification Kit (#K0702; Fermentas GmbH) aufgereinigt.

5.3.1.1.1 Übersicht PCR-Bedingungen

Primer	CasProm-F-34712024	5' - GTG CAT TAA CGG CAT CTG CA - 3'
	CasProm-R-2462603	5' - TGC AGA GGC AGC TGC TGA C - 3'

Ansatz Mastermix

Mastermix	Je Probe [µl]
H ₂ O ad 10µl	7,9
Puffer mit MgCl ₂	2,0
Q-Solution	4,0
Forwardprimer [10µM]	0,5
Reverseprimer [10µM]	0,5
dNTPs [2mM]	2,0
Taq	0,1

Programm

94°C	3'
94°C	45''
60°C	30''
72°C	90''
72°C	10'

x35

Amplikon 1296bp

5.3.1.1.2 Aufreinigung der PCR-Produkte

Um dNTPs, kurze Oligo-Fragmente und Enzyme von den jeweiligen PCR-Produkten effizient zu trennen, wurden diese mit dem GeneJET™ PCR Purification Kit (#K0702; Fermentas GmbH) nach Herstellerangaben aufgereinigt. Alle Zentrifugationsschritte erfolgten für 60 Sekunden bei 13.000g und die DNA wurde mit 20µl *Elution Buffer* von der Säule eluiert. Anschließend wurde die Konzentration der aufgereinigten DNA mittels 'Nanophotometer' bestimmt (siehe Abschnitt 5.2.3).

5.3.1.2 Sequenzierungsreaktion

Die Sequenzierung der Detektionsstichprobe wurde ausgelagert und ist von den Mitarbeitern des Labors der SRD - Scientific Research and Development GmbH (Bad Homburg) durchgeführt

worden. Hierzu wurden 16ng des aufgereinigten Amplifikates mit je 5pmol des Primers CasProm-F-34712024 oder CasProm-R-2462603 in 7µl Probenvolumen in ein PCR-Gefäß vorlegt, transportsicher verpackt und anschließend in das Labor der SRD gesendet. Die Cycle Sequencing Reaktion mit dem BigDye®Terminator v.3.1 Cycle Sequencing Kit (Life Technologies GmbH), sowie die anschließende Aufreinigung der DNA und die Kapillarelektrophorese auf dem 3130xL Genetic Analyzer wurde dort nach Herstellerangaben durchgeführt.

5.3.1.3 Auswertung der Rohdaten

Die Rohdaten aus der Sequenzanalyse wurden mit dem Programm ‚Seqscape‘ (Life Technologies GmbH) im Molekulargenetischen Labor der KJPPP bearbeitet und ausgewertet. Alle Sequenzen wurden hierbei mit einer von NCBI (National Center of Biotechnology) bezogenen genomischen Referenzsequenz (NT_007914.15|:6407832-6409832; Gen-Sequenz siehe Abschnitt 10.6) verglichen.

5.3.2 Auswertung der Sequenzierungsdaten

Da in den anschließenden Studien nur noch die Varianten genotypisiert werden sollten, die in der Detektionsstichprobe am häufigsten vorkamen (Genotypfrequenz > 1%) bzw. im Transmission-Disequilibrium-Test (TDT) eine Assoziation mit ASS aufwiesen (p-Wert < 0,05) erfolgte nach der Sequenzierung eine erste Auswertung der gewonnenen Daten. Ebenso wurde die Stichprobe auf vorliegende Mendelfehler überprüft. Für Details zur statistischen Auswertung siehe Abschnitt 5.6.

5.4 Analyse der Replikationsstichprobe

In der Replikationsstudie wurden die Einzelnukleotidpolymorphismen (SNPs) rs150447075, rs34712024, sowie der Short Tandem Repeat (STR) rs71781329 analysiert. Die Genotypisierung der Proben erfolgte im Molekulargenetischen Labor der KJPPP mittels RFLP. Ein Teil der Proben wurde am DKFZ in Heidelberg mittels Real-Time PCR analysiert. Die Genotypisierung der Variante rs71781329 wurde vollständig im Molekulargenetischen Labor der KJPPP in Frankfurt durchgeführt. Hierfür stellte das DKFZ Aliquots der DNA-Proben zur Verfügung.

5.4.1 Genotypisierung der Replikationsstichprobe mittels RFLP

Bei der Analyse von RFLPs (Restriktionsfragmentlängenpolymorphismus) macht man sich die Tatsache zu Nutze, dass der Basenaustausch eines SNPs zum Verlust bzw. Zugewinn einer Schnittstelle für spezifische Restriktionsendonukleasen führt.

Hierzu wurde der die zu analysierende Variante umfassende Abschnitt der extrahierten DNA zunächst mittels PCR vermehrt. Anschließend wurde das Amplikon mit einem speziell für die jeweilige Variante ermittelten Restriktionsenzym verdaut. Die Auftrennung der Verdauprodukte erfolgt mittels horizontaler oder vertikaler Gelelektrophorese. Je nach eingebauter Base können so unterschiedliche Fragmente erkannt werden.

Bei allen hier beschriebenen RFLPs erfolgte der Verdau von 4µl PCR-Produkt mit 1U des jeweiligen Enzyms in einem Reaktionsvolumen von 10µl. Die Inkubationszeit bei 37°C betrug 6 Stunden. Die Zugabe von Puffern und BSA erfolgte nach Herstellerangaben.

5.4.1.1 Genotypisierung der Variante rs150447075 (NG_007092.2:g.4595T>G)

Mit den unten angeführten Primern CasProm-F-34712024 und CasProm-R-34712024 wurde ein 409bp großes Fragment amplifiziert. Die PCR erfolgte nach Herstellerangaben mit der DreamTaq™ Polymerase (#EP0704; Fermentas GmbH). Das PCR-Produkt wurde anschließend mit dem Restriktionsenzym *Cac8I* (NEB GmbH) verdaut.

5.4.1.1.1 Übersicht PCR-Bedingungen

Primer	CasProm-F-34712024	5' - GTG CAT TAA CGG CAT CTG CA - 3'
	CasProm-R-34712024	5' - CCC TGG CCT CGT TTG AAA G - 3'

Ansatz Mastermix

Mastermix	Je Probe [µl]
H ₂ O ad 20µl	11,85
Puffer mit MgCl ₂	2,0
Forwardprimer [10µM]	0,5
Reverseprimer [10µM]	0,5
dNTPs [2mM]	2,0
Taq	0,15

Programm

95°C	3'
95°C	45"
64°C	30"
72°C	30"
72°C	10'

x30

Amplikon 409bp

5.4.1.1.2 Restriktionsverdau

Bei dem Verdau des PCR-Produktes mit dem Restriktionsenzym *Cac8I* (NEB GmbH) entsteht bei Vorliegen des Major-T-Allels ein Bandenmuster mit den Fragmentlängen 323, 85 und 1bp. Durch das G-Allel (minor) wird eine zusätzliche Schnittstelle für das Enzym an Position 307 eingefügt,

wodurch man ein Bandenmuster mit den Fragmentlängen 307, 85, 16 und 1bp erhält. Die Proben wurden nach dem Verdau vollständig auf ein 2%iges High-Resolution Agarosegel (2g Agarose High Resolution gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingssystem (AlphaMager HP, ProteinSimple) photographiert und mit der „AlphaView System Software“ (ProteinSimple) ausgewertet (Abbildung 3).

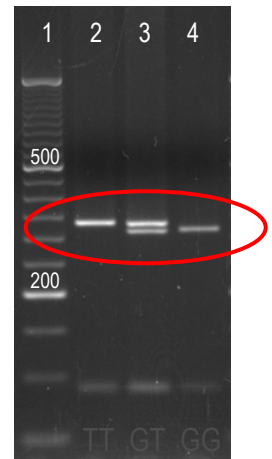


Abbildung 3: Gelbild nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-34712024 und Verdau des PCR-Produktes mit *Cac8I* (NEB GmbH); 1 = *O'RangeRuler™* 50bp DNA Ladder; ready-to-use (#SM0613; Fermentas GmbH); 2 = homozygot TT; 3 = heterozygot G/T; 4 = homozygot G/G.

5.4.1.2 Genotypisierung der Variante rs34712024 (NG_007092.2:g.4641A>G)

Mit den unten angeführten Primern CasProm-F-34712024 und CasProm-R-34712024 wurde ein 409bp großes Fragment amplifiziert. Die PCR erfolgte nach Herstellerangaben mit der DreamTaq™ Polymerase (#EP0704; Fermentas GmbH). Das PCR-Produkt wurde anschließend mit dem Restriktionsenzym *AluI* (Fermentas GmbH) verdaut.

5.4.1.2.1 Übersicht PCR-Bedingungen

Primer	CasProm-F-34712024	5' - GTG CAT TAA CGG CAT CTG CA - 3'
	CasProm-R-34712024	5' - CCC TGG CCT CGT TTG AAA G - 3'

Ansatz Mastermix

Mastermix	Je Probe [µl]
H ₂ O ad 20µl	11,85
Puffer mit MgCl ₂	2,0
Forwardprimer [10µM]	0,5
Reverseprimer [10µM]	0,5
dNTPs [2mM]	2,0
Taq	0,15

Programm

95°C	3'
95°C	45"
64°C	30"
72°C	30"
72°C	10'

x30

Amplikon 409bp

5.4.1.2.2 Restriktionsverdau

Bei dem Verdau des PCR-Produktes mit dem Restriktionsenzym *AluI* (Fermentas GmbH) entsteht bei Vorliegen des A-Allels (major) ein Bandenmuster mit den Fragmentlängen 219, 133 und 57bp.

Durch das G-Allel (minor) entfällt die Schnittstelle an Position 352, wodurch man ein Bandenmuster mit den Fragmentlängen 276 und 133bp erhält. Die Proben wurden nach dem Verdau vollständig auf ein 2%iges High-Resolution Agarosegel (2g Agarose High Resolution gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingsystem (Alphamager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 4).

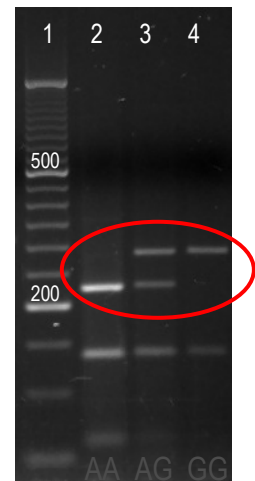


Abbildung 4: Gelbild nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-34712024 und Verdau des PCR-Produktes mit *AluI* (Fermentas GmbH); 1 = *O'RangeRuler™* 50bp DNA Ladder; ready-to-use (#SM0613; Fermentas GmbH); 2 = homozygot AA; 3 = heterozygot AG; 4 = homozygot GG.

5.4.1.3 Genotypisierung des STR rs71781329 (NG_007092.2: g.4856GCG(6_8))

Mit den Primern Caspr2-Prom-R2 und Caspr2-Prom-R2 wurde ein 319bp großes PCR-Produkt amplifiziert. Die PCR erfolgte nach Herstellerangaben mit der AmpliTaq® Gold DNA Polymerase (#4311820; Life Technologies GmbH) unter Zugabe von 0,8% DMSO. Das PCR-Produkt wurde anschließend mit dem Restriktionsenzym *HpaII* (Fermentas GmbH) verdaut.

5.4.1.3.1 Übersicht PCR-Bedingungen

Primer	Caspr2-Prom-F2	5' - CAG GGC AGA AGG GTT TTG ACA - 3'
	Caspr2-Prom-R2	5' - CTC ACA TGG AGA GCT TGT GTA TGT G - 3'

Ansatz Mastermix

Mastermix	Je Probe [µl]
H ₂ O ad 10µl	8,65
Puffer ohne MgCl ₂	2,0
MgCl ₂	1,6
DMSO	1,6
Forwardprimer [10µM]	0,5
Reverseprimer [10µM]	0,5
dNTPs [2mM]	2,0
Taq	0,15

Programm

95°C	3'
95°C	45"
58°C	30"
72°C	30"
72°C	10'

x30

Amplikon 319bp

5.4.1.3.2 Restriktionsverdau

Bei dem Verdau des PCR-Produktes mit *Hpa*II (Fermentas GmbH) entstehen bei Vorliegen von 6 GCG-Wiederholungen des Repeats Fragmente mit einer Länge von 139, 80, 49 und 45bp. Liegen in einem Allel sieben Wiederholungen des GCG-Motivs vor, verlängert sich das 80bp Fragment um 3bp auf 83bp. Dementsprechend liegt bei 8 Wiederholungen des Motivs ein 86bp langes Fragment vor.

5.4.1.3.3 Polyacrylamid-Gelelektrophorese (PAGE)

Wie die Agarose-Gelelektrophorese ist die Polyacrylamid-Gelelektrophorese (PAGE) ein Trennverfahren, bei dem Moleküle im elektrischen Spannungsfeld (35mA) nach ihrer Größe und Ladung aufgetrennt werden. Das Polyacrylamid, ein Polymer von Acrylamid, fungiert dabei als molekulares Sieb. Nach Zugabe von Ammoniumpersulfat (Radikalspender) und TEMED (Katalysator) erfolgt die Polymerisation von Acrylamid und Bis-Acrylamid durch eine radikalische Kettenreaktion. Für genauere Angaben zur Gelherstellung siehe Abschnitt 4.4.3.

Je 5 µl Verdauprodukt des STRs rs71781329 wurden auf einem 7,5%igen Polyacrylamidgel aufgetrennt. Um die aufgetrennte DNA in den Gelen sichtbar zu machen, wurde das Gel anschließend für 10 Minuten in einem Färbebad (0,2x GelRed in 0,1M NaCl) angefärbt. Alle Gele wurden für die Dokumentation und Auswertung mittels digitalem Imagingssystem (AlphaImager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 5).

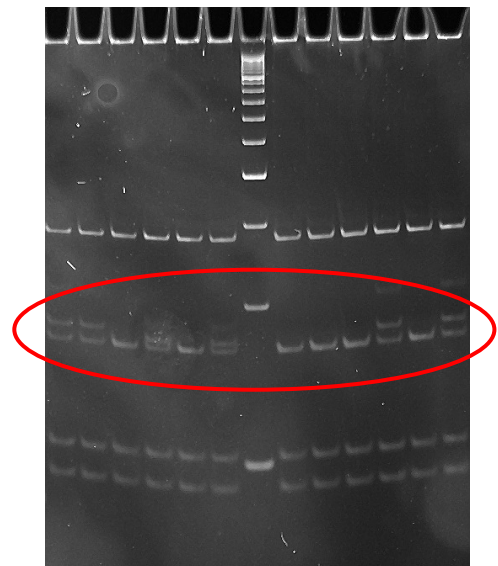


Abbildung 5: Gelbild nach Amplifikation mit Primerpaar Caspr2-Prom-F2 und Caspr2-Prom-R2 und Verdau des PCR-Produktes mit *Hpa*II (Fermentas GmbH); 7 = *O'RangeRuler™* 50bp DNA Ladder; ready-to-use (#SM0613; Fermentas GmbH); 3/5/8-10/12 = GCG(6_6); 4/6 = GCG(6_7); 1/2/11/13 = GCG(6_8).

5.4.2 Genotypisierung der Replikationsstichprobe mittels Real-Time PCR

Die Real-Time PCR beruht auf dem Prinzip der herkömmlichen PCR. Für die Genotypisierung werden allelspezifische Sonden eingesetzt, die mit fluoreszierenden Farbstoffen (FAM und VIC) markiert sind.

markiert sind. Diese binden zunächst an die komplementäre, einzelsträngige DNA und werden dann bei der Amplifikation mit eingebaut. Ein Anlagern der falschen Sonde ist kaum gegeben, da diese durch die energetisch stabilere Verbindung der exakt passenden Sonde verdrängt wird. Über das Verhältnis der beiden Fluoreszenzsignale erhält man dann die Information, welches der beiden Allele in der Probe vorliegt. Liegen zum Beispiel die Farbstoffe VIC und FAM bei der Messung in einem Verhältnis von 2:1 oder größer vor, trägt das Amplikon nur das Allel, welches sich auf der VIC markierten Sonde befindet (homozygot). Wenn das Verhältnis zwischen den Signalen ungefähr bei 1 liegt, ist die Probe heterozygot für die möglichen Merkmalsausprägungen.

Ein Teil der Genotypisierung der Varianten rs150447075 und rs34712024 wurde in der Replikationsstichprobe mittels Real-Time PCR am Deutschen Krebsforschungszentrum in Heidelberg durchgeführt. Die Analyse wurde von den dortigen Mitarbeitern mit dem speziell für diese Variante entwickelten Custom TaqMan® SNP Assays AH5IQWH (rs150447075) und AH39SP9 (rs34712024) auf dem 7900HT TaqMan System (Life Technologies GmbH) im 384well-Format durchgeführt. Das Probenvolumen lag bei 5µl.

5.5 Funktionelle Analyse der Varianten

5.5.1 In-silico Analyse möglicher Transkriptionsfaktorbindestellen in der *CNTNAP2*-Promotorregion

Die in-silico Analyse der Transkriptionsfaktorbindestellen in der *CNTNAP2*-Promotorregion erfolgte mit der Genomatrix Software MatInspector V2.6 (Quandt et al. 1995). Analysiert wurden die Sequenzen aller möglichen Allele der Varianten rs150447075, rs34712024 und rs71781329 \pm 50bp. Die ermittelten Ergebnisse der Majorallele wurden anschließend mit den Ergebnissen der jeweiligen Minorallele verglichen. Transkriptionsfaktoren, die in 3'-5'-Orientierung an den DNA-Strang binden, wurden als mögliche Regulatoren der *CNTNAP2*-Expression ausgeschlossen. Ein weiteres Ausschlusskriterium war, wenn die jeweilige Matrixsequenz die Variante nicht beinhaltet, bzw. der „Similarity Score“ unter 0,75 lag.

5.5.2 Luciferase Assay

Durch den Luciferase Assay ist es möglich, regulatorische DNA-Abschnitte auf ihre regulatorische Kapazität hin zu testen bzw. die Auswirkung von Promotorvarianten auf die Genregulation zu

untersuchen. Hierzu wird der zu untersuchende DNA-Abschnitt in einen sogenannten „Basic“-Reporter-Vektor vor das Luciferase-Gen kloniert und anschließend in ein adäquates Zellmodell eingebracht. Die bei dieser Methode eingesetzte Firefly Luciferase ist ein 61kDa schweres, monomeres Enzym, welches unter Anwesenheit des Co-substrates ATP-Mg²⁺ die Oxidation von Luciferin katalysiert. Die dabei entstehenden energiereichen, instabilen Substanzen zerfallen unter Abgabe von Licht (Biolumineszenz), dessen Intensität luminometrisch bestimmt werden kann.

Ob die Varianten rs150447075, rs34712024 und rs71781329 einen Effekt auf die Promotoraktivität haben, wurde im Rahmen dieser Studie an den humanen Zelllinien HEK293T und SH-SY5Y (Tabelle 2) getestet, welche freundlicherweise von Frau Prof. Dr. Fulda aus dem Institut für Experimentelle Tumorforschung in der Pädiatrie (Frankfurt/Main) zur Verfügung gestellt wurden. Die Quantifizierung der Luciferase-Aktivität erfolgte bei den HEK293T Zellen 24 Stunden nach der Transfektion, bei den SH-SY5Y Zellen erst nach 48 Stunden. Des Weiteren erfolgte in den SH-SY5Y Zellen eine Messung nach 72 und 216 Stunden nach Induktion der Differenzierung.

Tabelle 2: Auflistung der verwendeten humanen Zelllinien.

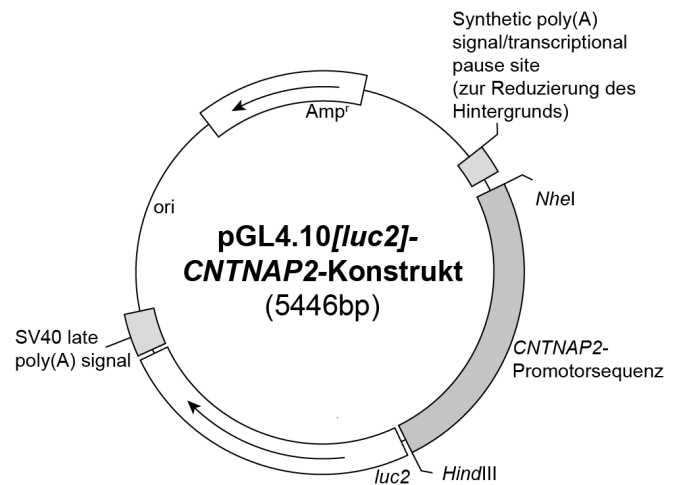
Zelllinien	Ursprung	ATCC®-Nummer
HEK 293T	Humane Embryonale Nierenzellen	CRL-1573™
SH-SY5Y	Humane Neuroblastomzelllinie	CRL-2266™

5.5.2.1 Vektorkonstruktion

Die in-silico Vektorkonstruktion erfolgte mit dem Programm Geneious 5.5.3. An die 5'-Enden der Klonierungsprimer AC_1.2 und AC_1.3 wurden Erkennungssequenzen für die Restriktionsenzyme *NheI* und *HindIII* angehängt. Der *sense*-Primer enthält somit die Erkennungssequenz für das Restriktionsenzym *NheI* und der *antisense*-Primer die Erkennungssequenz für das Restriktionsenzym *HindIII*. Das PCR Produkt wurde unter Anwendung einer hochspezifischen Polymerase amplifiziert und mittels bluntend-Ligation in das pGEM® Vektorsystem I (Promega GmbH) zwischenkloniert. Positive Klone wurden durch RFLP und direkter Sequenzierung validiert. Durch einen Doppelverdau mit *NheI* und *HindIII* kann einerseits der Zielvektor pGL4.10[*luc2*] linearisiert werden und andererseits das Insert für die Umklonierung aus dem pGEM®-*CNTNAP2*-Vektorkonstrukt herausgeschnitten werden. Die dabei entstehenden Überhänge sorgen dafür, dass

die Ligation des Insert in den Zielvektor in 5'-3'-Orientierung erfolgt. Das fertige pGL4.10-CNTNAP2-Konstrukt ist in Abbildung 6 dargestellt.

Abbildung 6: Darstellung des pGL4.10[luc2]-CNTNAP2-Vektorkonstrukts (nach: www.promega.com). Die inserierte CNTNAP2-Promotorsequenz ist in grün dargestellt. Abkürzungen: Amp^r, Ampicillinresistenzgen; luc2, synthetisches Fireflyluciferasegen; ori, Replikationsursprung; SV40, Simian-Virus 40.



5.5.2.2 Klonierung

5.5.2.2.1 Amplifikation der CNTNAP2-Promotorregion für die Klonierung

Mit den Primern AC_1.2 und AC_1.3 wurde ein 1248bp großes Fragment amplifiziert, welches die CNTNAP2-Promotorregion beinhaltet. Die PCR erfolgte nach Herstellerangaben mit der PRECISOR High-Fidelity DNA Polymerase (#1706-250-BL; BioCat GmbH).

5.5.2.2.1.1 Übersicht PCR-Bedingungen

Primer	AC_1.2	5' - g [^] ct agc ATT CCA GGA TTA GGC AAG CCT CCA - 3'
	AC_1.3	5' - a [^] ag ctt TTG TGT ATG TGG GGA GCG CGC A - 3'

Ansatz Mastermix

Mastermix	Je Probe [μl]
H ₂ O ad 25μl	14,5
5x GC-Buffer	5,0
Forwardprimer [10μM]	1,0
Reverseprimer [10μM]	1,0
dNTPs [2mM]	3,0
Taq	0,5

Programm

98°C	2'
98°C	30''
69°C	30''
72°C	45''
72°C	7'

x30

Amplikon 1248bp

5.5.2.2.1.2 Aufreinigung der PCR-Produkte

Um dNTPs, kurze Oligo-Fragmente oder Enzyme von den jeweiligen PCR-Produkten effizient zu trennen, wurden diese mit dem GeneJET™ PCR Purification Kit (#K0702; Fermentas GmbH) nach Herstellerangaben aufgereinigt (siehe Abschnitt 5.3.1.1.2) und deren Endkonzentration mittels Implen ‚Nanophotometer‘ bestimmt (siehe Abschnitt 5.2.3)

5.5.2.2.2 Ligation der Insert-DNA mit dem pGEM®-T Easy Vektor

Die Ligation des PCR-Produktes mit dem pGEM®-T Easy Vektor (pGEM®-T Easy Vector System I; #A1360; Promega GmbH) erfolgte in einem 10µl Ansatz nach Herstellerangaben. Die Vektor- und Insert-DNA wurden im molaren Verhältnis 1:10 eingesetzt. Um die Ausbeute bei der Blunt-End Ligation zu erhöhen erfolgte der Inkubationsschritt über Nacht im Eiswasserbad.

5.5.2.2.3 Transformation des pGEM®-CNTNAP2-Konstrukts in die chemisch kompetenten *E.coli* Zellen

Von dem Ligationsansatz wurden am folgenden Tag 5µl in die Transformation eingesetzt. Nach Zugabe der chemisch kompetenten *E.coli* Zellen (DH5α™ Competent Cells; #18265-017; Life Technologies GmbH) wurde der Ansatz für 1 Stunde auf Eis inkubiert. Vor der Zugabe von 950µl LB-Medium erfolgte der Hitzeschock der Zellen für 50 Sekunden bei 42°C. Abschließend wurde der fertige Ansatz für 90min bei 37°C und 150rpm inkubiert. Je 100µl der transformierten Kulturen wurden auf die mit 80µl der X-Gal-Lösung (40mg/ml) behandelten LB/Ampicillin-Platten ausplattiert und über Nacht bei 37°C inkubiert. Die durch die Blau-Weiß-Reaktion vorselektierten positiven weißen Klone wurden zusätzlich mittels Kolonie-PCR kontrolliert.

5.5.2.2.4 Kolonie-PCR

Die Kolonie-PCR dient der Kontrolle, ob die Insertion eines bestimmten DNA-Abschnittes in einen Vektor erfolgreich war. Für die Kolonie-PCR wurden die entsprechenden Kolonien mit einer Pipettenspitze gepickt und diese für 2 Minuten in den bereits in das PCR-Gefäß vorgelegten Mastermix gestellt. Die Spitze wurde anschließend aus dem Gefäß entfernt und verworfen. Die PCR erfolgte nach Herstellerangaben mit der PRECISOR High-Fidelity DNA Polymerase (#1706-250-BL; BioCat GmbH) und den Primern AC_1.2 und AC_1.3. Das Amplifikat wurde je nach vorliegender Variante mit dem entsprechenden Restriktionsenzym verdaut.

5.5.2.2.4.1 Übersicht PCR-Bedingungen

Primer	AC_1.2	5' - g [^] ct agc ATT CCA GGA TTA GGC AAG CCT CCA - 3'
	AC_1.3	5' - a [^] ag ctt TTG TGT ATG TGG GGA GCG CGC A - 3'

Ansatz Mastermix

Mastermix	Je Probe [μl]
H ₂ O ad 25μl	14,5
5x GC-Buffer	5,0
Forwardprimer [10μM]	1,0
Reverseprimer [10μM]	1,0
dNTPs [2mM]	3,0
Taq	0,5

Programm

98°C	2'
98°C	30''
69°C	30''
72°C	45''
72°C	7'

x30

Amplikon 1248bp

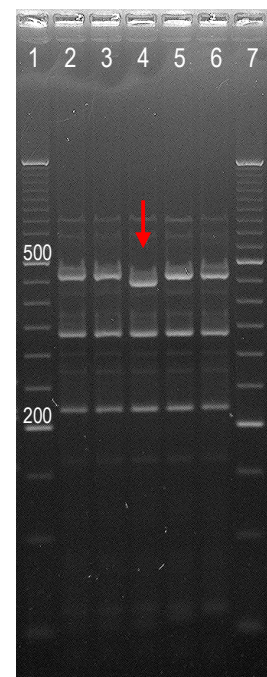
5.5.2.2.4.2 Verdau des Produktes aus der Kolonie-PCR

Für den Verdau von 1μl des Produktes aus der Kolonie-PCR wurde 1U des jeweiligen Enzyms in ein Reaktionsvolumen von 10μl eingesetzt. Die Inkubationszeit bei 37°C betrug 6 Stunden. Die Zugabe von Puffern und BSA erfolgte nach Herstellerangaben.

Verdau des Kolonie-PCR Produktes mit *Cac8I*

Bei dem Verdau des Produktes aus der Kolonie-PCR mit dem Restriktionsenzym *Cac8I* (NEB GmbH) entsteht bei Vorliegen von Allel rs150447075T (major) ein Bandenmuster mit den Fragmentlängen 455, 334, 218, 56, 40, 23, 20, 18 und 16bp. Durch das Allel rs150447075G (minor) wird eine zusätzliche Schnittstelle für das Enzym an Position 829 eingefügt, wodurch man ein Bandenmuster mit den Fragmentlängen 439, 334, 218, 56, 40, 23, 20, 18 und 16bp erhält. Die Probe wurden nach dem Verdau vollständig auf ein 2%iges High-Resolution Agarosegel (2g Agarose High Resolution gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingsystem (AlphaImager HP, ProteinSimple) photographiert und mit der 'AlphaView System Software' (ProteinSimple) ausgewertet (Abbildung 7).

Abbildung 7: Gelbild nach Amplifikation mit Primerpaar AC_1.2 und AC_1.3 und Verdau des PCR-Produktes mit *Cac8I* (NEB GmbH); 1/7 = *O'RangeRuler™* 50bp DNA Ladder; ready-to-use (#SM0613; Fermentas GmbH); 2/3/5/6 = rs150447075T; 4 = rs150447075G.



Verdau des Kolonie-PCR Produktes mit *AluI*

Bei dem Verdau des Produktes aus der Kolonie-PCR mit dem Restriktionsenzym *AluI* (Fermentas GmbH) entsteht bei Vorliegen von Allel rs34712024A (major) ein Bandenmuster mit den Fragmentlängen 380, 255, 219, 210, 152, 13 und 7bp. Durch das Allel rs34712024G (minor) entfällt die Schnittstelle an Position 874, wodurch man ein Bandenmuster mit den Fragmentlängen 429, 380, 255, 152, 13 und 7 erhält. Die Probe werden nach dem Verdau vollständig auf ein 2%iges High-Resolution Agarosegel (2g Agarose High Resolution gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingsystem (AlphaMager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 8).

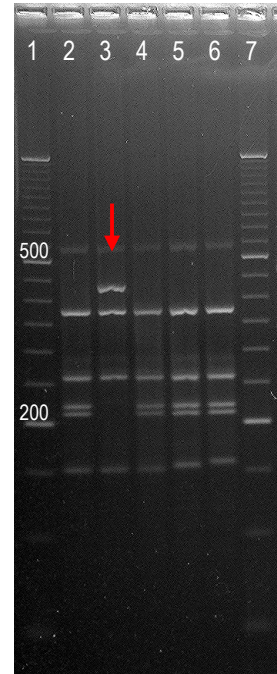


Abbildung 8: Gelbild nach Amplifikation mit Primerpaar AC_1.2 und AC_1.3 und Verdau des PCR-Produktes mit *AluI* (Fermentas GmbH); 1/7 = *O'RangeRuler™* 50bp DNA Ladder; ready-to-use (#SM0613; Fermentas GmbH); 2/4-6 = rs34712024A; 3 = rs34712024G.

Verdau des Kolonie-PCR Produktes mit *HpaII*

Bei dem Verdau des Produktes aus der Kolonie-PCR mit dem Restriktionsenzym *HpaII* (Fermentas GmbH) entstehen bei Vorliegen von Allel rs71781329GCG[6] Fragmente mit einer Länge von 643, 181, 142, 93, 80, 49, 30, 8 und 4bp. Liegen sieben Wiederholungen des GCG-Motivs vor, verlängert sich das 80bp Fragment um 3bp auf 83bp. Dementsprechend liegt bei 8 Wiederholungen des Motivs ein 86bp langes Fragment vor.

Je 5µl des Verdauproduktes wurden auf einem 7,5%igen Polyacrylamidgel aufgetrennt. Um die aufgetrennte DNA in den Gelen sichtbar zu machen, wurde das Gel anschließend für 10 Minuten in einem Färbegrad (0,2x GelRed in 0,1M NaCl) angefärbt. Alle Gele wurden für die Dokumentation und Auswertung mittels digitalem Imagingsystem (AlphaMager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 9).

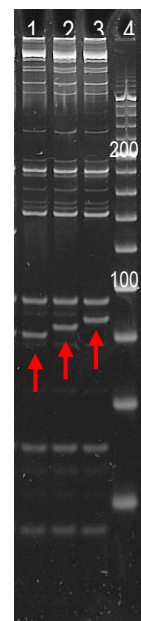


Abbildung 9: Gelbild nach Amplifikation mit Primerpaar AC_1.2 und AC_1.3 und Verdau des PCR-Produktes mit *HpaII* (Fermentas GmbH); 1 = rs71781329GCG[6]; 2 = rs71781329GCG[7]; 3 = rs71781329GCG[8]; 4 = *O'RangeRuler™* 20bp DNA Ladder; ready-to-use (#SM1323; Fermentas GmbH).

5.5.2.2.5 Vermehrung der pGEM®-CNTNAP2-Konstrukte: Miniprep

Um das pGEM®-CNTNAP2-Konstrukt zu vermehren wurden 3,6ml des nach Herstellerangaben hergestellten LB-Mediums mit 0,4ml 10x HMFM und 4µl Ampicillin versetzt. Anschließend wurde die entsprechende Kolonie gepickt und über Nacht bei 37°C und 150rpm in dem Flüssignährmedium inkubiert.

Für die Präparation der Plasmid-DNA mit dem GeneJET™ Plasmid Miniprep Kit (#K0503; Fermentas GmbH) wurden 1,5ml Bakterienkultur in ein frisches Reaktionsgefäß gegeben und für 2 Minuten bei Raumtemperatur und 6800g abzentrifugiert. Der Überstand wurde vollständig entfernt und die anschließende Präparation der DNA erfolgte nach Herstellerangaben.

5.5.2.2.6 Kontrollverdau des pGEM®-CNTNAP2-Konstrukts

Wenn die CNTNAP2-Promotorregion in den pGEM®-T Easy Vektor eingebaut wurde erhält man bei dem Verdau von 1µl des pGEM®-CNTNAP2-Konstruktes bei 37°C mit 1U des Restriktionsenzym *DraI* (Fermentas GmbH) ein Bandenmuster von 2283, 1266 und 692bp. Das gesamte Probenvolumen von 10µl wurde nach der sechsständigen Inkubation auf ein 0,8%iges Agarosegel (0,8g SeaKem® LE Agarose gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingsystem (AlphaMager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 10).

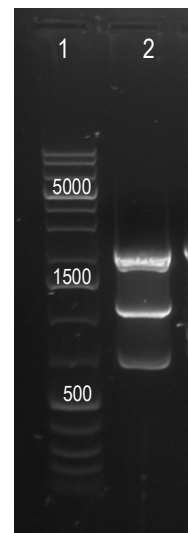


Abbildung 10: Gelbild nach Verdau von 1µl des pGEM®-CNTNAP2-Konstruktes mit *DraI* (Fermentas GmbH);
1 = O'GeneRuler™ 1kb Plus DNA Ladder; ready-to-use (#SM1334; Fermentas GmbH); 2 = pGEM®-CNTNAP2-Konstrukt.

5.5.2.3 Umklonierung des Inserts in Zielvektor pGL4.10[*luc2*]

Für die Umklonierung in den Zielvektor wurde das Insert zunächst mit Hilfe von Restriktionsendonukleasen herausgeschnitten, bzw. der Zielvektor linearisiert und das entsprechende DNA-Fragment mit Hilfe eines präparativen Agarosegels isoliert. Abschließend wurde das Insert mit dem Zielvektor pGL4.10[*luc2*] ligiert und das Konstrukt zur Vermehrung in chemisch kompetente Zellen transformiert. Das Konstrukt wurde nach jedem Schritt durch RFLP und abschließend durch Sequenzierung kontrolliert

5.5.2.3.1 Herausschneiden des Inserts / Linearisierung des Zielvektors pGL4.10[*luc2*]

Das Herausschneiden des Inserts, bzw. die Linearisierung des Vektors pGL4.10[*luc2*] (#E6651; Promega GmbH) erfolgte in einem Doppelverdau mit den Restriktionsendonukleasen *NheI* und *HindIII* (NEB GmbH). Hierzu wurden 5µl der DNA aus der Plasmidpräparation (Miniprep) und je 1U der Enzyme in einem 10µl Reaktionsansatz für 6 Stunden bei 37°C verdaut.

5.5.2.3.2 Isolierung der DNA aus dem präparativen Agarosegel

Um ausschließlich das Insert bzw. den linearisierten Vektor zu erhalten wurde die gesamte Probe aus dem Verdau auf ein 0,8%iges (0,8g SeaKem® LE Agarose gelöst in 100ml 1x TTE-Puffer), präparatives Agarosegel aufgetragen, die Bande gesuchter Größe ausgeschnitten und die DNA mit dem GeneJET™ PCR Purification Kit (Fermentas GmbH) aus dem Gel aufgereinigt. Hierzu wurde das Gelstück im Verhältnis 1:1 (Gewicht:Volumen) mit *Binding Buffer* versetzt und bei 50°C inkubiert, bis das Gel vollständig gelöst war. Alle weiteren Schritte wurden nach Herstellerangaben durchgeführt: Die Zentrifugationsschritte erfolgten für 60 Sekunden bei 13.000g und die DNA wurde mit 30µl Aqua ad iniectabilia (B. Braun Melsungen AG) von der Säule eluiert. Anschließend wurde die Konzentration der DNA mittels Implen ‚Nanophotometer‘ bestimmt (siehe Abschnitt 5.2.3)

5.5.2.3.3 Ligation der Insert-DNA mit dem Zielvektor pGL4.10[*luc2*]

Die Ligation des Inserts mit dem Zielvektor pGL4.10[*luc2*] (#E6651; Promega GmbH) wurde unter Anwendung einer T4 DNA Ligase (#EL0016; Fermentas GmbH) nach Herstellerangaben durchgeführt. Die Insert-DNA wurde im molaren Verhältnis 5:1 eingesetzt (Insert:Vektor).

5.5.2.3.4 Transformation des pGL4.10[*luc2*]-*CNTNAP2*-Konstrukts in die chemisch kompetenten *E.coli* Zellen

Von dem Ligationsansatz wurden 5µl nach Standardbedingungen in die chemisch kompetenten *E.coli* Zellen (DH5α™ Competent Cells; #18265-017; Life Technologies GmbH) transformiert (siehe Abschnitt 5.5.2.2.3). Von den transformierten Kulturen wurden je 150µl auf LB/Ampicillin-Platten ausplattiert und über Nacht bei 37°C inkubiert. Die gewachsenen Kolonien wurden mittels Kolonie-PCR und anschließendem Verdau darauf kontrolliert, ob sie das Zielkonstrukt tragen (siehe Abschnitt 5.5.2.2.4 und 5.5.2.2.4.2).

5.5.2.3.5 Vermehrung der pGL4.10[*luc2*]-*CNTNAP2*-Konstrukte: Midiprep

Um genügend Material für die weiteren Experimente zu gewinnen wurden die positiven Klone vermehrt. Hierzu wurden 45ml des nach Herstellerangaben hergestellten LB-Mediums mit 5ml 10x HMFM und 50µl Ampicillin versetzt, die entsprechenden Kolonien gepickt und über Nacht bei 37°C und 150rpm in dem Flüssignährmedium inkubiert.

Vor der Präparation der Plasmid-DNA mit dem GeneJET™ Plasmid Midiprep Kit (#K0482; Fermentas GmbH) wurden 50ml Bakterienkultur in ein frisches Reaktionsgefäß gegeben und für 10 Minuten bei Raumtemperatur und 5000g abzentrifugiert. Der Überstand wurde vollständig entfernt und die anschließende Präparation erfolgte nach Herstellerangaben. Die Elution der Plasmid-DNA von der Säule erfolgt mit Aqua ad iniectabilia (B. Braun Melsungen AG).

5.5.2.3.6 Kontrollverdau des pGL4.10[*luc2*]-*CNTNAP2*-Konstrukts

Um zu kontrollieren, ob das Insert in der richtigen Orientierung im pGL4.10[*luc2*]-*CNTNAP2*-Konstrukt vorliegt, wurde die isolierte Plasmid-DNA mit dem Restriktionsenzym *DraI* (Fermentas GmbH) verdaut. Bei dem Verdau von 1µl der Plasmid-DNA bei 37°C mit 1U des Enzyms erhält man ein Bandenmuster von 3106, 1265, 1056 und 19bp. Das gesamte Probenvolumen von 10µl wurde nach der sechsständigen Inkubation auf ein 0,8%iges Agarosegel (0,8g SeaKem® LE Agarose gelöst in 100ml 1x TTE-Puffer) aufgetragen und mittels digitalem Imagingsystem (AlphaMager HP, ProteinSimple) photographiert und mit der ‚AlphaView System Software‘ (ProteinSimple) ausgewertet (Abbildung 11).

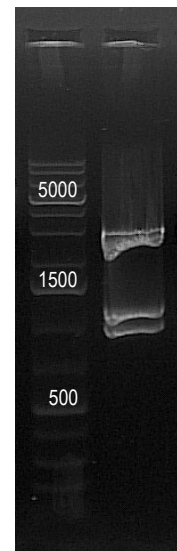


Abbildung 11: Gelbild nach Verdau von 1µl des pGL4.10[*luc2*]-*CNTNAP2*-Konstruktes mit *DraI* (Fermentas GmbH); 1 = *O'GeneRuler™* 1kb Plus DNA Ladder; ready-to-use (#SM1334; Fermentas GmbH); 2 = pGL4.10[*luc2*]-*CNTNAP2*-Konstrukt.

5.5.2.3.7 Kontrolle des Inserts durch direkte Sequenzierung

Um ungewollt eingefügte Mutationen auszuschließen, wurden die Plasmide vor der Transfektion sequenziert. Die Sequenzkontrolle des Inserts wurde ausgelagert und ist somit von den Mitarbeitern des Labors der SRD - Scientific Research and Development GmbH (Bad Homburg) durchgeführt worden. Für die Sequenzierung wurden 300ng des Plasmids mit 5pmol des Primers AC_1.2 bzw. AC_1.3 in ein PCR-Gefäß vorlegt (7µl Probenvolumen), transportsicher verpackt und anschließend

in das Labor der SRD gesendet. Die Cycle Sequencing Reaktion mit dem BigDye®Terminator v.3.1 Cycle Sequencing Kit (Life Technologies GmbH) sowie die anschließende Aufreinigung der DNA und die Kapillarelektrophorese mittels 3130xL Genetic Analyzer (Life Technologies GmbH) wurden dort nach Herstellerangaben durchgeführt. Die Sequenzen wurden im Molekulargenetischen Labor der KJPPP mit dem Programm ‚Seqscape‘ (Life Technologies GmbH) bearbeitet und ausgewertet. Referenzsequenz: NT_007914.15|:6407832-6409832 (Gen-Sequenz: siehe Abschnitt 10.6).

5.5.2.4 Zellkultur

5.5.2.4.1 Kultivierung der Zellen

Die Kultivierung der Zelllinien HEK293T und SH-SY5Y erfolgte in dem mit 10% FCS, 1% Pyruvat, 100U/ml Penicillin und 100µg/ml Streptomycin angereicherten DMEM+GlutaMAX Medium (Life Technologies GmbH) bei 37°C in 5%iger CO₂ Atmosphäre. Je nach Typ und Konfluenz der Zellen wurde das Nährmedium alle 3-5 Tage abgezogen, die Zellen mit Dulbecco's 1x PBS ohne Mg²⁺ und Ca²⁺ (PAA Laboratories GmbH) gewaschen und durch die Zugabe von 0,5-1ml Trypsin EDTA 0,05% (Life Technologies GmbH) vom Flaschenboden abgelöst. Um die Reaktion des Trypsin EDTA zu stoppen wurde mindestens die 5fache Menge serumhaltiges Medium dazu gegeben und die Zellen im Verhältnis 1:2 – 1:20 in neue Kulturflaschen überführt.

Für die verschiedenen experimentellen Ansätze wurden die Zellen in einer Neubauerzählkammer gezählt, verdünnt und in benötigter Dichte in die entsprechenden Zellkulturplatten ausgesät.

5.5.2.4.2 Zellen einfrieren

Das Nährmedium wurde zunächst von den Zellen abgezogen. Anschließend wurden die Zellen mit Dulbecco's 1x PBS (PAA Laboratories GmbH) gewaschen, durch die Zugabe von 0,5-1ml Trypsin EDTA 0,05% (Life Technologies GmbH) vom Flaschenboden abgelöst, mit 1-3ml Standardmedium in eine Zentrifugenröhre (15ml) überführt und für 5 Minuten bei 700g abzentrifugiert. Der Überstand wurde vorsichtig abgesaugt, das Pellet in 2-4ml Freezemedium resuspendiert und je nach Zelldichte auf 3-4 Cryoröhrchen verteilt. Anschließend wurden die Röhrchen bei -80°C in einer Kryo-Einfrierbox (Kühlrate: 1°C/Stunde) eingefroren.

5.5.2.4.3 Zellen auftauen

Da das Einfriermedium für die Zellen toxisch ist, muss das Auftauen so schnell wie möglich von statten gehen. Daher wurde das Cryoröhrchen mit den Zellen im Wasserbad (37°C) schwenkend

aufgetaut. Nach Zugabe von ca. 2ml vorwärmtem Standardmedium (37°C) wurden die Zellen in eine Zentrifugenröhre (15ml) überführt und für 5 Minuten bei 700g abzentrifugiert. Der Überstand wurde vorsichtig abgesaugt und das Pellet in 2ml Standardmedium resuspendiert. Je nach Zelllinie wurden die Zellen auf 1-2 Zellkulturflaschen verteilt und mit Standardmedium aufgefüllt (T25 ad 4-6ml; T75 ad 12-15ml).

5.5.2.5 Transfektionseffizienztest

Da für die Neuroblastomzelllinie SH-SY5Y die für die HEK293T-Zelllinie etablierte Transfektionsmethode nicht verwendet werden konnte, wurde zunächst die Transfektionseffizienz verschiedener Reagenzien getestet (Tabelle). Unter Berücksichtigung der jeweiligen Herstellerangaben wurde geschaut, bei welchem Reagenz und unter welchen Bedingungen die Zellen der SH-SY5Y Zelllinie am effizientesten mit dem Vektor pmaxGFP® (Lonza GmbH) transfiziert werden (Tabelle 4). Bei erfolgter Transfektion der Zellen, exprimieren diese das grünfluoreszierende Protein, welches wiederum fluoreszenzmikroskopisch nachgewiesen werden kann. Folgende Reagenzien wurden getestet:

Tabelle 3: Name und Hersteller der beim Transfektionseffizienz-Test getesteten Reagenzien.

Reagenz	Hersteller
Attractene (#1051561)	Qiagen
Metafectene® Easy+ (#T090-0.1)	Biontex Laboratories GmbH
Metafectene® Pro (#T040-0.2)	Biontex Laboratories GmbH
Fugene® HD (#E2313)	Promega GmbH
Xfect™ (#631316)	Takara Bio Europe SAS
TransIT®-2020 (#MIR 5404)	Mirus Bio LLC
TurboFect (#R0539)	Fermentas GmbH

5.5.2.5.1 Durchführung des Transfektionseffizienztests

Die SH-SY5Y Zellen wurden in einer 24-well Zellkulturplatte mit einer Dichte von 1×10^5 Zellen/ml und 500µl/Well ausgesät. Direkt vor der Transfektion wurde das Medium abgezogen und pro Well 250µl frisches Medium zugegeben. Die Transfektion erfolgte unter folgenden Bedingungen:

Tabelle 4: Testbedingungen der verschiedenen Transfektionsreagenzien.

Reagenz	Menge DNA [μ g]	Verhältnis DNA [μ g] / Reagenz [μ l]
Attractene	0,2 / 0,4 / 0,6	1:1,5 / 1:3,75 / 1:7,5
Metafectene® Easy+	0,6 / 1,3 / 1,9	1:1
Metafectene® Pro	0,5 / 1 / 1,5	1:1 / 1:2 / 1:4 / 1:6
Fugene® HD	0,5 / 1 / 1,5	1:1 / 1:2 / 1:3 / 1:4
Xfect™	0,5 / 1 / 2	0,3 μ l Polymer/ μ g DNA
TransIT®-2020	0,5 / 1 / 1,5	1:1 / 1:2 / 1:3 / 1:4
TurboFect	0,5 / 1 / 1,5	1:1 / 1:2 / 1:3 / 1:4

Des Weiteren wurde geschaut, ob sich die Transkriptionsrate erhöht, wenn der Ansatz in Dulbecco's 1x PBS (PAA Laboratories GmbH) erfolgt, beziehungsweise die Transfektion bei einer höheren Zelldichte (1×10^5 Zellen/ml, 2×10^5 Zellen/ml, 4×10^5 Zellen/ml) durchgeführt wird. Die Auswertung der erstellten fluoreszenzmikroskopischen Aufnahmen erfolgte halbautomatisch.

5.5.2.5.2 Auswertung der fluoreszenzmikroskopischen Aufnahmen

Von einem frei gewählten, mäßig bewachsenen Bereich aus der Mitte jedes Wells wurde sowohl eine lichtmikroskopische, als auch eine fluoreszenzmikroskopische Aufnahme mit 20facher Vergrößerung erstellt. Die Zellzahl beider Aufnahmen wurde halbautomatisch unter Anwendung des Programms ImageJ 1.46r (<http://rsb.info.nih.gov/ij/>) bestimmt. Aus der ermittelten Zellzahl der licht- und der fluoreszenzmikroskopischen Aufnahme konnte dann der prozentuale Anteil an transfizierten Zellen in dem jeweiligen Bildausschnitt ermittelt werden. Folgende Einstellungen wurden im Programm vorgenommen:

5.5.2.5.2.1 Einstellungen für die Zellzahlbestimmung in der lichtmikroskopischen Aufnahme (Hellfeld)

- 1) Image - Adjust - Brightness/Contrast → auto
- 2) Process - Noise – Remove Outliers...
 - Radius = 5px
 - Threshold = 2
 - Which outliers = Dark
- 3) Process - Find Maxima...
 - Noise tolerance = 50

Output type = Count

→ Exclude edge maxima

→ Light background

→ Preview point selection: Über die Auswahl "Preview point selection" kann man sich anzeigen lassen, welche Zellen unter der eingestellten "Noise tolerance" gezählt werden. Die Toleranz sollte so eingestellt werden, dass die markierten Zellen der Zellzahl entsprechen.

5.5.2.5.2.2 Einstellungen für die Zellzahlbestimmung in der fluoreszenzmikroskopischen Aufnahme (Dunkelfeld)

1) Process - Noise – Remove Outliers...

Radius = 15px

Threshold = 5

Which outliers = Bright

2) Process - Find Maxima...

Noise tolerance = 10

Output type = Count

→ Exclude edge maxima

→ Preview point selection: Über die Auswahl "Preview point selection" kann man sich anzeigen lassen, welche Zellen unter der eingestellten "Noise tolerance" gezählt werden. Die Toleranz sollte so eingestellt werden, dass die markierten Zellen der Zellzahl entsprechen.

Die finale Transfektion der Zelllinien für den Luciferaseassay erfolgte unter den ermittelten optimalen Transfektionsbedingungen.

5.5.2.5.3 Differenzierung der Neuroblastomzellen SH-SY5Y

In Anwesenheit von Retinsäure und BDNF (Brain-Derived-Neurotrophic-Factor) differenzieren die Zellen der SH-SY5Y Zelllinie zu reifen Neuronen (Encinas et al. 2000). Um durch die Varianten hervorgerufene Veränderungen in der Promotoraktivität während der Differenzierungsphase nachzuweisen, wurde die Luciferaseaktivität in den Zellen zu verschiedenen Entwicklungszeitpunkten gemessen. Die Messpunkte ergeben sich aus der Versorgung der Zellen

mit dem Neurobasal-A-Medium und hBDNF/RA und entsprechen dem Zeitraum: 24 Stunden nach Mediumwechsel (Tabelle 5).

Der Zeitpunkt 1 liegt vor der ersten Versorgung der Zellen mit dem Neurobasal-A-Medium und hBDNF/RA. Es liegen daher ausschließlich undifferenzierte Zellen vor und entsprechen dem Zeitraum: 48 Stunden nach der Transfektion. Vor dem Wechsel vom Standardmedium zum Neurobasal-A-Medium mit hBDNF/RA werden die Zellen mit Dulbecco's 1x PBS (PAA Laboratories GmbH) gewaschen.

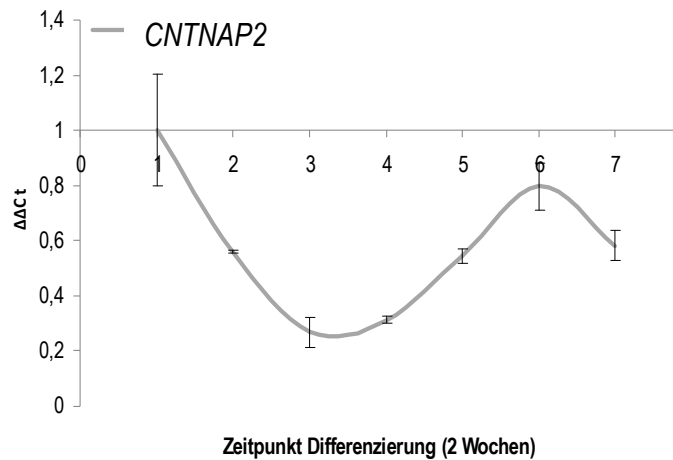
Tabelle 5: Tabellarische Darstellung der zeitlichen Durchführung der Differenzierungsstudie.

Tag	Zeitpunkt	Versuch #1	Versuch #2	Versuch #3
1		Aussaat der Zellen in Standardmedium	Aussaat der Zellen in Standardmedium	Aussaat der Zellen in Standardmedium
2		Mediumwechsel / Transfektion in Standardmedium	Mediumwechsel / Transfektion in Standardmedium	Mediumwechsel / Transfektion in Standardmedium
3				
4	#1	Messung #1	Mediumwechsel hBDNF/RA-Medium	Mediumwechsel hBDNF/RA-Medium
5	#2			
6			Mediumwechsel hBDNF/RA-Medium	Mediumwechsel hBDNF/RA-Medium
7	#3		Messung #3	
8				Mediumwechsel hBDNF/RA-Medium
9	#4			
10				Mediumwechsel hBDNF/RA-Medium
11	#5			
12				Mediumwechsel hBDNF/RA-Medium
13	#6			Messung #6

Die Auswahl der Zeitpunkte an denen die Luciferaseaktivitätsmessung in den differenzierenden Zellen durchgeführt wurde, erfolgte aufgrund einer Proteinexpressionsstudie von Frau Denise Haslinger (Msc). In dieser konnte sie zeigen, dass die Expression von *CNTNAP2* in den ersten Tagen der Zelldifferenzierung runterreguliert wird. Zum Zeitpunkt 3, also 72 Stunden nach Induktion der Zelldifferenzierung, hat dieser Effekt seinen tiefsten Punkt erreicht. Zum Ende der Differenzierung (Zeitpunkt 6 → 216 Stunden) normalisiert sich die *CNTNAP2*-Expressionsrate wieder (Abbildung 12). Die Kontrolle des Zelldifferenzierungsstatus erfolgte über die Expression der

neuronalen Marker TAU (Hochregulierung) und CDC2 (Zellteilungsmarker; Runterregulierung). Die Daten von Frau Haslinger werden hier nicht gezeigt.

Abbildung 12: Expressionsanalyse von *CNTNAP2* in SH-SY5Y Zelllinie während der Differenzierungsphase; die Ergebnisse und Abbildung wurden aufgrund persönlicher Kommunikation von Frau Denise Haslinger (Msc) zur Verfügung gestellt (nicht publizierte Daten).



5.5.2.5.4 Transfektion der pGL4.10[luc2]-*CNTNAP2*-Konstrukte

Die ersten Transfektionen der pGL4.10[luc2]-*CNTNAP2*-Konstrukte in die HEK293T Zellen wurden in Kooperation mit dem Institut für Experimentelle Tumorforschung in der Pädiatrie (Frankfurt/Main) durchgeführt. Diese erfolgten nach Herstellerangaben mit dem X-tremeGENE HP DNA Transfection Reagent (Roche Diagnostics GmbH). Das Verhältnis von Reagenz zu DNA betrug 3:1.

Die im Molekulargenetischen Labor der KJPPP durchgeführten Transfektionen wurden mit dem Transfektionsreagenz Metafectene® Pro (#T040-2.0; Biontex Laboratories GmbH) nach den ermittelten Transfektionsbedingungen durchgeführt (siehe Abschnitt 6.4.2.1.4). Das Verhältnis von Reagenz zu DNA betrug 4:1.

Transfiziert wurde der leere Vektor pGL4.10[luc2] (#E6651; Promega GmbH), beziehungsweise die unterschiedlichen pGL4.10[luc2]-*CNTNAP2*-Konstrukte. Als Kontrolle für die Transfektionseffizienz erfolgte eine Cotransfektion des Vektors pGL4.74[hRLuc/TK] (#E6921; Promega GmbH) im Verhältnis 100:1 (pGL4.10[luc2]: pGL4.74[hRLuc/TK]).

5.5.2.5.4.1 Transfizierte Vektoren / Vektorkonstrukt

pGL4.10[luc2]	Leervektor (Normalisierung)	LV
FAUT-0850-005/pGL4.10[luc2]	Majorallelkonstrukt (Referenz)	WT
FAUT-0858-001/pGL4.10[luc2]	rs34712024G	AG
FAUT-0811-005/pGL4.10[luc2]	rs150447075G	GT
FAUT-0370-003/pGL4.10[luc2]	rs71781329GCG[7]	Ins1
HAUT-0361-002/pGL4.10[luc2]	rs71781329[8]	Ins2
pGL4.74[hRLuc/TK]	Kontrollvektor für Transfektionseffizienz	

5.5.2.5.4.2 Durchführung der Transfektion

Die HEK293T Zellen wurden in einer 12-well Zellkulturplatte mit einer Dichte von 2×10^5 Zellen/ml und 1ml/Well, die SH-SY5Y Zellen mit einer Dichte von 4×10^5 Zellen/ml und 1ml/Well ausgesät. Für die Differenzierungsstudie wurden die SH-SY5Y Zellen ebenfalls in einer 12-well Zellkulturplatte ausgesät, jedoch mit einer Dichte von 2×10^5 Zellen/ml und 1ml/Well. In allen drei Fällen wurde nach 24 Stunden das Medium abgezogen und pro Well 1ml frisches Medium zugegeben. Die Transfektion erfolgte unter folgenden Bedingungen:

Transfektionsansatz (pro Well)

Ansatz 1

50µl Dulbecco's 1x PBS (PAA Laboratories GmbH)

1µg pGL4.10[luc2] (#E6651; Promega GmbH) bzw. pGL4.10[luc2]-CNTNAP2-Konstrukte

0,01µg pGL4.74[hRLuc/TK] (#E6921; Promega GmbH)

Ansatz 2

50µl Dulbecco's 1x PBS (PAA Laboratories GmbH)

4µl Metafectene® Pro (Biontex Laboratories GmbH)

Ansatz 1 wurde vorsichtig zu Ansatz 2 gegeben und durch einmaliges auf- und abpipettieren gemischt. Nach einer Inkubationszeit von 15 Minuten wurden 100µl des Transfektionsansatzes tropfenweise zu den Zellen eines Wells gegeben, die Platte vorsichtig geschüttelt und für 24 Stunden (HEK293T) beziehungsweise 48 Stunden (SH-SY5Y) bei 37°C in 5%iger CO₂ Atmosphäre inkubiert.

5.5.2.6 Quantifizierung der Luciferaseaktivität

5.5.2.6.1 Quantifizierung der Luciferaseaktivität in der HEK293T-Zelllinie und den undifferenzierten Zellen der SH-SY5Y Zelllinie

Die Quantifizierung der Firefly Luciferase Aktivität (*luc+*) in der HEK293T-Zelllinie und den undifferenzierten Zellen der SH-SY5Y Zelllinie erfolgte nach Standardprotokoll mit dem Dual-Luciferase® Reporter Assay System (#E1910; Promega GmbH). In jeder Zelllinie wurde die Messung in drei biologischen Replikaten durchgeführt. Innerhalb der biologischen Replikate lagen von jeder Variante vier technische Replikate vor. Anschließend erfolgte die Auswertung der Messergebnisse.

5.5.2.6.2 Messung der Luciferaseaktivität in der SH-SY5Y Zelllinie

Die Quantifizierung der Firefly Luciferase Aktivität (*luc+*) zu unterschiedlichen Differenzierungszeitpunkten der SH-SY5Y Zelllinie erfolgte mit dem Dual-Luciferase® Reporter Assay System (Promega GmbH). Da die Zellen für die Differenzierung mit einer geringeren Dichte ausgesät wurden, mussten die Zellen aus 2 Wells vereint werden. Nur so war gewährleistet, dass bei der Lichtmessung genug Signal entsteht. Hierzu wurden die Zellen nach Entfernen des Mediums und Waschen mit Dulbecco's 1x PBS ohne Mg^{2+} und Ca^{2+} (PAA Laboratories GmbH) mit je 100µl Trypsin EDTA 0,05% (Life Technologies GmbH) abtrypsiniert und der gesamte Inhalt von 2 Wells in ein frisches 1,5ml Reaktionsgefäß überführt. Die Trypsinierungsreaktion wurde mit 800µl Standardmedium gestoppt und die Zellen für 5 Minuten bei 700g abzentrifugiert. Nachdem das Medium vorsichtig vom Zellpellet entfernt und die Zellen ein weiteres Mal mit Dulbecco's 1x PBS (PAA Laboratories GmbH) gewaschen wurden, erfolgte die Zellyse. Hierzu wurden die Zellen in 100µl *Passive Lysis Buffer* (1x) resuspendiert und für 20 Minuten bei 150rpm im Schüttelinkubator inkubiert. Von dem Zellysat wurden jeweils 40µl in die Messung eingesetzt, welche nach Herstellerangaben erfolgte. Für die statistische Auswertung siehe Abschnitt 5.6.

5.6 Statistische Auswertung

5.6.1 Deskriptive Statistik

Bei der deskriptiven Statistik werden die erhobenen Daten in Form von Zahlen, Tabellen und Grafiken dargestellt. Die Berechnung der relativen Häufigkeiten der Allelfrequenzen in der Detektionsstichprobe erfolgte mit dem Programm Haploview Version 4.1 (Barrett et al. 2005). Die deskriptive Statistik der Replikationsstichprobe und des Gesamtsets wurde mit dem Programm IBM SPSS Statistics 20.0 durchgeführt.

5.6.2 Poweranalyse

Um Aufschluss darüber zu bekommen, wie hoch die Chance ist, dass in einer Studie ein bestimmter Effekt wirklich nachgewiesen wird, können Poweranalysen durchgeführt werden. Die berechnete Power ist abhängig von der Stichprobengröße, der Effektstärke und der Streuung. Die Poweranalysen erfolgten mit dem Programm Quanto Version 1.2.4 (Gauderman W).

5.6.3 Genetische Analyse

5.6.3.1 Transmission Disequilibrium Test (TDT)

Mit Hilfe des TDT kann eine Aussage darüber gemacht werden, ob bzw. welches Allel überzufällig häufig von einem heterozygoten Elternteil an ein erkranktes Kind weitergegeben wird. Üblicherweise werden hierbei ausschließlich vollständige Kernfamilien, also Familien, bei denen der Genotyp des Kindes und beider Elternteile ermittelt werden konnte, in die Berechnung eingeschlossen. Hier wurde die Berechnung des TDT mit UNPHASED 3.1.4 (Dudbridge 2008) durchgeführt. da dieses Programm bei nicht vollständigen Kernfamilien den Genotyp des fehlenden Elternteils schätzt, wodurch sich die Power erhöht.

5.6.3.2 Überprüfung auf Mendelfehler

Bei der Überprüfung auf Mendelfehler wird aufgrund der Familienstruktur geprüft, ob die ermittelten Genotypen mit der Mendelschen Vererbung vereinbar sind. Der Nachweis möglicher Mendelfehler erfolgte mit dem Programm Haploview Version 4.1 (Barrett et al. 2005). Proben mit einem nachgewiesenen Mendelfehler wurden aus der Stichprobe entfernt.

5.6.3.3 Nachweis von Imprinting- oder geschlechtsspezifischen Effekten

Der vorliegende Phänotyp kann zusätzlich durch Imprinting- oder geschlechtsspezifische Effekte beeinflusst werden. Unter Imprinting versteht man das Phänomen, dass Chromosomen von väterlicher und mütterlicher Herkunft funktionell ungleich und somit ursächlich für einen bestimmten Phänotyp sein können. Bei einem geschlechtsspezifischen Effekt hingegen wird der vorliegende Phänotyp durch das Geschlecht des Patienten beeinflusst. Der Nachweis möglicher Imprinting- oder geschlechtsspezifischer Effekte erfolgte mit UNPHASED 3.1.4 (Dudbridge 2008).

5.6.3.4 Regressionsanalyse

Mit Hilfe von Regressionsanalysen kann man den Zusammenhang zwischen einer abhängigen und einer oder mehreren unabhängigen Variablen feststellen. Da die Sprach-Items A9 (Alter in Monaten, in dem die ersten Worte gesprochen wurden) und A10 (Alter in Monaten, in dem die ersten Sätze gesprochen wurden) des ADI-R Tests nicht normalverteilt vorliegen (Kolmogorov-Smirnov $p < 0,001$), wurde eine nicht-parametrische Regressionsanalyse durchgeführt, um den Effekt der dichotomisierten Genotypen auf die Sprach-Items A9 und A10 zu untersuchen. In diesem Fall wurde ein dominantes Modell vorausgesetzt und die Items wurden auf Alter und IQ korrigiert. Die Analyse

wurde in SAS Version 9.3 (SAS Institute Inc.) mit dem dafür verfügbaren Macro ‚npar‘ (<http://www.ams.med.uni-goettingen.de/Projekte/makros/index.html>) mit freundlicher Unterstützung von Prof. Dr. Freitag durchgeführt.

5.6.4 Auswertung Luciferase Assay

Für die Auswertung der Ergebnisse der drei biologischen Replikate wurden zunächst alle Messungen auf die Transfektionseffizienz korrigiert. Möglich war das durch die Cotransfektion des Renilla Luciferase exprimierenden Vektors pGL4.74[*hRluc*/TK] (#E6921; Promega GmbH). Anschließend wurden die Mittelwerte der jeweiligen technischen Replikate gebildet und die Firefly Luciferase Aktivität (*luc*+) aller transfizierten Vektoren auf den Leervektor pGL4.10[*luc2*] (#E6651; Promega GmbH) normalisiert und gegen das ausschließlich Majorallele tragende Konstrukt verglichen. Mittels ANOVA für Messwiederholungen wurde abschließend geschaut, ob die Konstrukte signifikante Unterschiede in der Luciferase-Aktivität (*luc*+) zeigen. Die Auswertung der Ergebnisse erfolgte mit Microsoft® Office Excel 2003 und dem Programm IBM SPSS Statistics 20.0.

6 Ergebnisse

6.1 Beschreibung der Stichproben

Die deskriptive Statistik zeigte, dass sich die Detektionsstichprobe und die Replikationsstichprobe in Bezug auf Alter bei Diagnose, IQ, Geschlecht und Diagnose nicht signifikant unterschieden. In der Gesamtstichprobe lag das mittlere Alter der Patienten zum Zeitpunkt der Diagnose bei 10,75 Jahren. Der durchschnittliche IQ-Wert lag bei 88,51. 86% der Patienten waren männlich und 14% waren weiblich was einem Geschlechterverhältnis von 6,1:1 (männlich:weiblich) entspricht. Im Vergleich zu dem aus der Literatur hervorgehenden Geschlechterverhältnis von 3-4:1 (Volkmar et al. 2004) ist das Geschlechterverhältnis der vorliegenden Stichprobe leicht zugunsten der männlichen Probanden verschoben. Des Weiteren wurde bei 87,5% der Patienten eine strikte Autismusdiagnose gestellt, welche den High (IQ \geq 70) und den Low Functioning (IQ < 70) Autismus umfasst. Bei 12,5% der Patienten wurde eine Spektrumsstörung diagnostiziert, welche das Asperger-Syndrom und den atypischen Autismus (PDD-NOS) einschließt. Die Ergebnisse der deskriptiven Statistik sind in Tabelle 6 aufgelistet.

Tabelle 6: Deskriptive Statistik der Detektions-, Replikations- und Gesamtstichprobe.

	N	Alter in Jahren (SD)	N	IQ (SD)	N	Geschlecht (Frequenz)	N	Diagnose (Frequenz)
Detektion	216	10,353 (5,443)	207	89,13 (25,744)	236	m (0,869) w (0,131)	236	A (0,856) S (0,144)
Replikation	353	10,997 (6,475)	317	88,10 (23,176)	356	m (0,857) w (0,143)	356	A (0,888) S (0,112)
p-Wert		U-test: 0,725		U-Test: 0,407		Chi²: 0,717		Chi²: 0,256
Gesamtset	569	10,75 (6,11)	524	88,51 (24,20)	592	m (0,861) w (0,139)	592	A (0,875) S (0,125)

Abkürzungen: A, Autismus; Chi², Chi²-Test; m, männlich; U-test, Mann-Whitney-U-Test; S, Spektrumsstörung; SD, Standardabweichung; w, weiblich.

6.2 Analyse der Detektionsstichprobe

Um sowohl bekannte, als auch bisher nicht beschriebene Varianten in der *CNTNAP2*-Promotorregion nachzuweisen, wurde diese Region in 236 Familien mittels direkter Sequenzierung analysiert. Die Auswertung der Sequenzdaten erfolgte mit dem Programm ‚SeqScape‘ (Life Technologies GmbH).

6.2.1 Sequenzierung der *CNTNAP2*-Promotorregion

In der analysierten Region liegen zehn Varianten, die bereits in verschiedenen Datenbanken (dbSNP, HapMap und UCSC; siehe Abschnitt 10.2) beschrieben wurden. Von diesen zehn Varianten kamen drei (rs150447075, rs34712024 und rs71781329) in den untersuchten 236 Familien der Detektionsstichprobe vor. Rs13225855, rs13225866, rs13225872, rs13226652, rs13225963, rs13225979 und rs13226843 konnten nicht nachgewiesen werden. Zusätzlich wurden sechs bisher nicht beschriebene Varianten in dieser Studie entdeckt (Tabelle 7).

6.2.1.1 Nachweis bereits beschriebener Varianten

6.2.1.1.1 Polymorphismus rs150447075

In 27 der 667 Personen der Detektionsstichprobe konnte mindestens ein Minor-G-Allel des Einzelnukleotidpolymorphismus rs150447075T>G nachgewiesen werden (Abbildung 13). Dies entspricht einer berechneten Minorallelfrequenz (MAF) von 0,023 (Tabelle 7), was sich mit der in Datenbanken beschriebenen Allelfrequenz der Gesamtbevölkerung deckt (MAF 0,022; dbSNP: www.ncbi.nlm.nih.gov/sites/SNP). Die Genotypfrequenz in den Patienten beträgt für GT und GG kombiniert 4,6%.

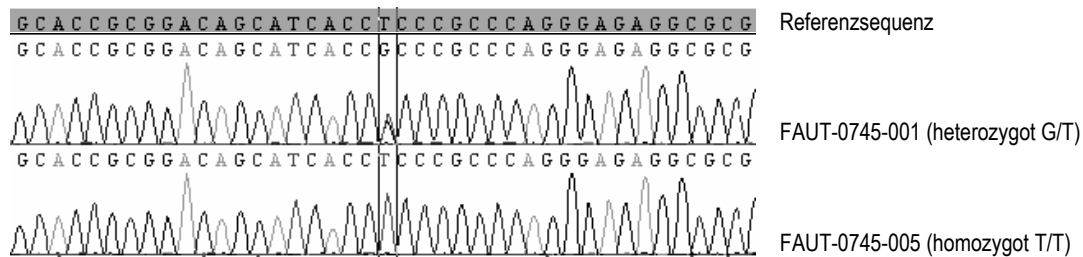


Abbildung 13: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position des SNP rs150447075.

6.2.1.1.2 Polymorphismus rs34712024

In 23 der 667 Personen der Detektionsstichprobe konnte mindestens ein Minor-G-Allel des Einzelnukleotidpolymorphismus rs34712024A>G nachgewiesen werden (Abbildung 14). Dies entspricht einer berechneten MAF von 0,011 (Tabelle 7). In der Gesamtbevölkerung kommt dieser Polymorphismus mit einer Allelfrequenz von 0,029 vor (<http://www.ncbi.nlm.nih.gov/sites/SNP>). Die Genotypfrequenz in den Patienten beträgt für AG und GG kombiniert 2,1%.

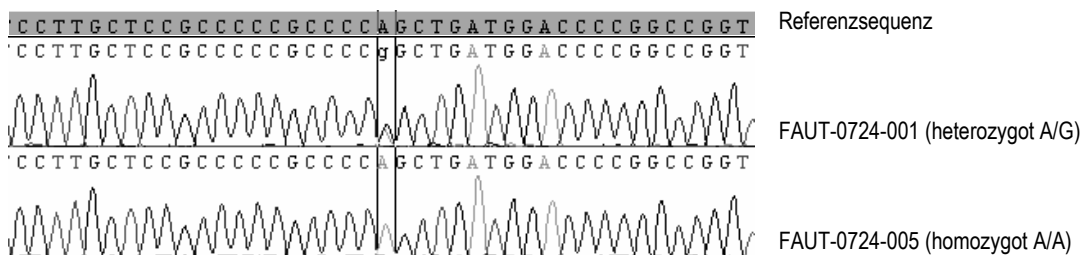


Abbildung 14: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position des SNP rs34712024.

6.2.1.1.3 Short Tandem Repeat rs71781329

Der Short Tandem Repeat (STR) rs71781329 liegt in der Stichprobe in drei Varianten vor. In zwei DNA-Proben der 667 Personen umfassenden Detektionsstichprobe konnte eine 7fache Wiederholung des GCG-Mikrosatelliten nachgewiesen werden (MAF = 0,002). In vier DNA-Proben wurde der 8fache Tripletteinschub GCG[8] nachgewiesen (MAF = 0,004) (Abbildung 15). Für beide Varianten zusammen entspricht dies einer Minorallelfrequenz von 0,006. Wie häufig die Variante in der Gesamtbevölkerung vorkommt ist nicht bekannt (<http://www.ncbi.nlm.nih.gov/sites/SNP>). Die Genotypfrequenz in den Patienten beträgt für GCG[7] und GCG[8] kombiniert 1,3%.

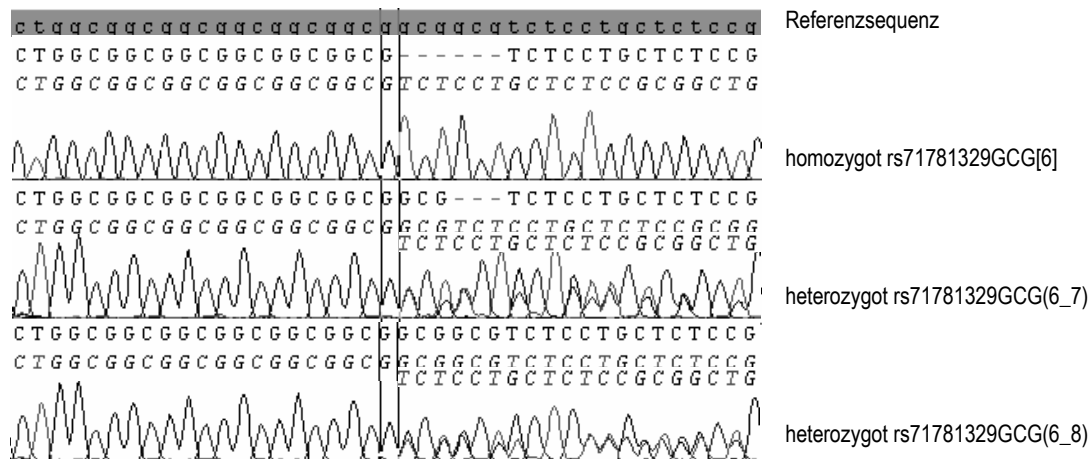


Abbildung 15: Ausschnitt aus dem Elektropherogramm der homozygoten Variante rs71781329GCG[6] und der heterozygoten Varianten rs71781329GCG(6_7) und rs71781329GCG(6_8) nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Base vor dem STR.

6.2.1.2 Nachweis bisher nicht beschriebener Varianten

In einzelnen Proben konnten sechs bisher nicht beschriebene Varianten (M1-M6) nachgewiesen werden.

6.2.1.2.1 M1: Deletion an Position 145813046 (GRCh37)

Die heterozygote Deletion (C→Del) konnte nur in einer nicht betroffenen Halbschwester eines Indexpatienten nachgewiesen werden, nicht aber in einem der Patienten (Abbildung 16). Diese Variante kam einmal in den analysierten 667 Proben vor.

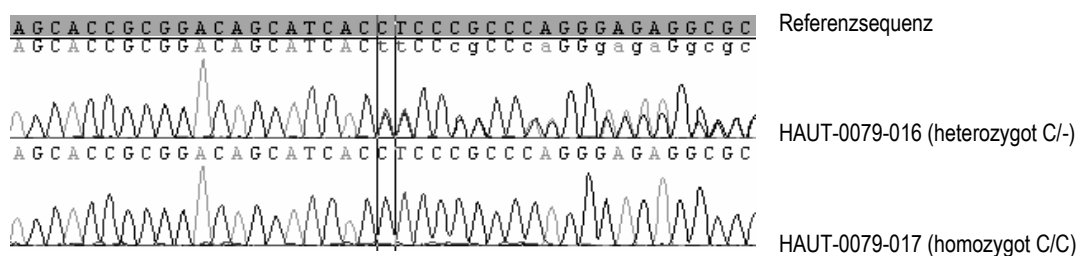


Abbildung 16: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M1.

6.2.1.2.2 M2: Mutation an Position 145813238 (GRCh37)

Der heterozygote Basenaustausch (G→A) wurde in einem Fall durch den Vater an das erkrankte Kind weitervererbt (Abbildung 17). Er kam somit insgesamt zweimal in den analysierten 667 Proben vor.

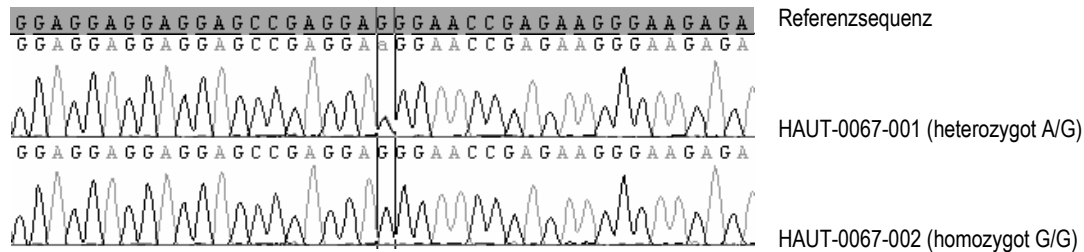


Abbildung 17: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M2.

6.2.1.2.3 M3: Mutation an Position 145813410 (GRCh37)

Der heterozygote Basenaustausch (G→C) konnten nur in einer Mutter nachgewiesen werden, nicht aber in einem der Patienten (Abbildung 18). Er kam nur einmal in den analysierten 667 Proben vor.

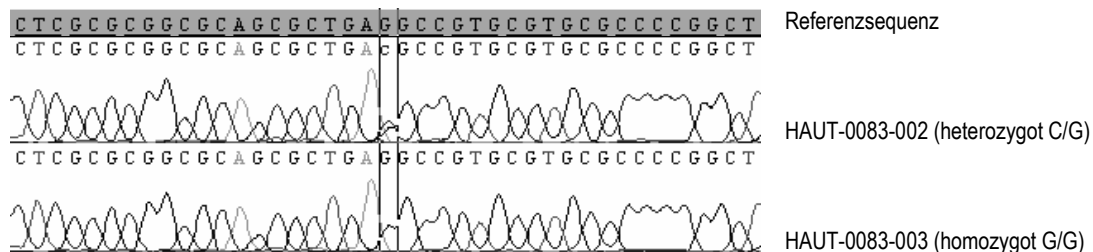


Abbildung 18: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M3.

6.2.1.2.4 M4: Mutation an Position 145813427 (GRCh37)

Der heterozygote Basenaustausch (G→C) wurde in einem Fall durch den Vater an das erkrankte Kind weitervererbt. Die Mutation kam zusätzlich bei einem weiteren Vater, einem nicht betroffenen Geschwisterkind (paternal vererbt) und einer Mutter, also insgesamt fünfmal in den analysierten 667 Proben vor (Abbildung 19).

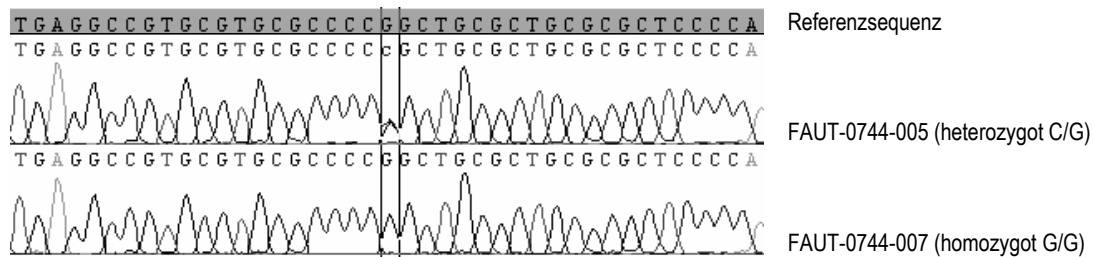


Abbildung 19: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M4.

6.2.1.2.5 M5: Mutation an Position 145813497 (GRCh37)

Der heterozygote Basenaustausch (G→A) wurde in einem Fall durch den Vater an das erkrankte Kind weitervererbt (Abbildung 20). Die Mutation kam somit insgesamt zweimal in den analysierten 667 Proben vor.

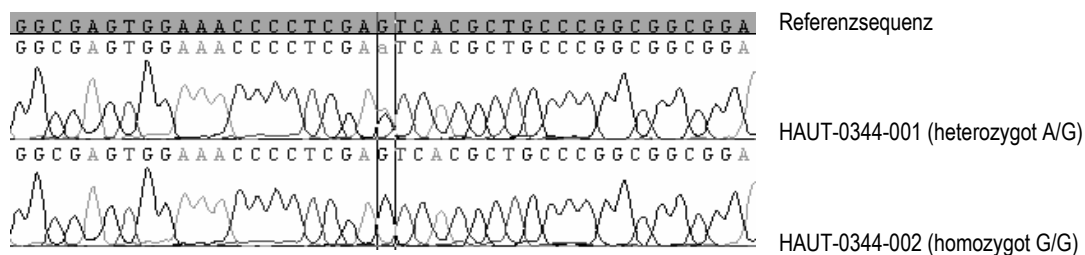


Abbildung 20: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M5.

6.2.1.2.6 M6: Mutation an Position 145813559 (GRCh37)

Der heterozygote Basenaustausch (C→G) wurde innerhalb einer einzigen Familie durch die Mutter an das erkrankte Kind, seine beiden nicht betroffenen Brüder und seine Halbschwester weitervererbt (Abbildung 21). Die Variante kam somit fünfmal in den analysierten 667 Proben vor.

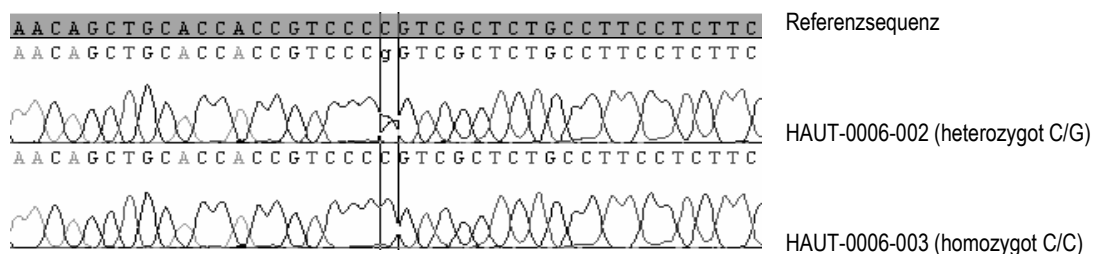


Abbildung 21: Ausschnitt aus dem Elektropherogramm der *CNTNAP2*-Promotorregion nach Amplifikation mit Primerpaar CasProm-F-34712024 und CasProm-R-2462603; Referenzsequenz = bekannte Sequenz aus NCBI Datenbank (NT_007914.15 | 6407832-6409832). Die Markierung entspricht der Position der Mutation M6.

Da die Genotypfrequenz der bisher nicht beschriebenen Varianten unter 1% liegt, wurden sie von den weiteren Berechnungen und Analysen ausgeschlossen. Die SNPs rs150447075 (NG_007092.2:g.4595T>G) und rs34712024 (NG_007092.2:g.4641A>G), als auch der trimere Short Tandem Repeat (STR) rs71781329 (NG_007092.2: g.4856GCG(6_8)) erreichten eine Genotypfrequenz über 1% und wurden daher in die weiteren Berechnungen und Analysen eingeschlossen. Die Ergebnisse der Sequenzanalyse sind detailliert in Tabelle 7 aufgelistet.

Tabelle 7: Übersicht der in der Detektionsstichprobe nachgewiesenen Varianten.

	SNP	Allele Major / Minor	N Index (236)	Vererbt von	MAF ^a	Position (GRCh37)
Bekannt	rs150447075	T/G	11	3 paternal; 5 maternal; 3 NV	G = 0,023	145813047
	rs34712024	A/G	5	1 paternal; 1 maternal; 3 NV	G = 0,011	145813093
	rs71781329	GCG[6]/GCG[7]	1	1 paternal	GCG[7] = 0,002	145813308
	kombiniert	GCG[6]/GCG[8]	2	1 paternal; 1 maternal	GCG[8] = 0,004	
		GCG[6]/GCG(7_8)	3	2 paternal; 1 maternal	GCG(7_8) = 0,006	
	rs13225855		NN			145813233
	rs13225866		NN			145813252
	rs13225872		NN			145813261
	rs13226652		NN			145813270
	rs13225963		NN			145813304
	rs13225979		NN			145813374
	rs13226843		NN			145813386
Neu	M1	C/-		Nur in Halbschwester nachgewiesen		145813046
	M2	G/A	1	1 paternal		145813238
	M3	G/C		Nur in Mutter nachgewiesen		145813410
	M4	G/C	1	1 paternal		145813427
	M5	G/A	1	1 paternal		145813497
	M6	C/G	1	1 maternal		145813559

^aMinor-Allelfrequenz der Varianten bei den erkrankten Kindern der Detektionsstichprobe. *Abkürzungen:* GRCh37, Genome Reference Consortium human genome (build 37); M, Mutation; MAF, Minorallelfrequenz; NN, Nicht nachweisbar; NV, Nicht vorhanden; SNP, Einzelnukleotidpolymorphismus.

6.3 Analyse der Proben aus der Replikationsstichprobe und statistische Auswertung der Genotypisierungsdaten

Zur Vergrößerung der Stichprobe erfolgte die Genotypisierung der drei häufigsten Varianten (rs150447075, rs34712024 und rs71781329) in weiteren 356 Familien. Anschließend wurden die Daten aus der Detektionsstichprobe und der Replikationsstichprobe statistisch ausgewertet.

6.3.1 Genotypisierung

Insgesamt war es möglich 99,73% (rs150447075 und rs34712024) und 98,37% (rs71781329) der Replikationsstichprobe zu genotypisieren (Tabelle 8).

6.3.2 Statistische Auswertung der Genotypisierungsdaten

Mit den Genotypisierungsdaten der drei häufigsten Varianten (rs150447075, rs34712024 und rs71781329) wurde ein Transmission Disequilibrium Test (TDT) durchgeführt. Eine Haplotypenanalyse entfiel jedoch, da nur bei einer einzigen Person das gleichzeitige Vorkommen der Minorallele der Varianten rs150447075 und rs34712024 nachgewiesen werden konnte.

6.3.2.1 Pre-test Poweranalyse

In der Poweranalyse wurde berechnet, welche Effekte in den einzelnen Stichproben mit einer Power von mindestens 80% nachgewiesen werden können, unter der Annahme der in der Sequenzierung detektierten Minorallelfrequenzen der Varianten rs150447075, rs34712024 und rs71781329, einer zweiseitigen Signifikanz von 0,05, einem dominanten Modell und einem Populationsrisiko von 1% (Tabelle 8).

Die Pre-test Analyse zeigte, dass bei den vorliegenden MAFs < 2% nur OR>2,90 für die Risikoallele detektiert werden können. Die Berechnung der Detektions- und der Replikationsstichprobe erfolgte trotz der geringen Power zunächst getrennt von einander. Um aber letztendlich eine ausreichende Power zu erhalten wurden die beiden Samples kombiniert und gemeinsam ausgewertet. Hierbei lag die unter den oben genannten Annahmen detektierbare Effektgrößen je nach MAF zwischen OR>2,06 und OR>3,44 für die Risikoallele. Die Ergebnisse der Poweranalyse für die oben gelisteten Annahmen und einer Power von 80% sind in Tabelle 8 detailliert aufgelistet.

Tabelle 8: Poweranalyse der Varianten rs150447075, rs34712024 und rs71781329 unter der Annahme einer zweiseitigen Signifikanz von 0,05, einem dominanten Modell und einem Populationsrisiko von 1%.

	MAF ^a	Detektionsstichprobe (N = 236 Familien)	Replikationsstichprobe (N = 356 Familien)	Gesamtstichprobe (N = 592 Familien)
rs150447075T/G	0,020	0,345 > OR > 2,900	0,410 > OR > 2,440	0,485 > OR > 2,060
rs34712024A/G	0,017	0,323 > OR > 3,100	0,387 > OR > 2,600	0,463 > OR > 2,160
rs71781329GCG[6]/GCG(7_8)	0,005	0,174 > OR > 5,740	0,223 > OR > 4,480	0,291 > OR > 3,440

^aNachgewiesene Minorallelfrequenz (MAF) der Detektionsstichprobe. Abkürzungen: MAF, Minorallelfrequenz; OR, Odds-Ratio

6.3.2.2 Assoziationsstudie

Für das Minor-G-Allel des SNP **rs150447075** konnte weder in der Detektionsstichprobe (OR = 1,000; CI95 = 0,322 – 3,101; p-Wert = 1,000) noch in der Replikationsstichprobe (OR = 0,888; CI95 = 0,453 – 1,744; p-Wert = 0,732) bzw. der Gesamtstichprobe (OR = 0,917; CI95 = 0,514 – 1,635; p-Wert = 0,768) eine Assoziation mit ASS nachgewiesen werden.

Für das Minor-G-Allel der Variante **rs34712024** hingegen ergaben die Berechnungen in der Detektionsstichprobe eine nominale Assoziation mit ASS (OR = 0,200; CI95 = 0,044 – 0,913; p-Wert = 0,016), die in der Replikationsstichprobe jedoch nicht repliziert werden konnte (OR = 0,583; CI95 = 0,230 – 1,482; p-Wert = 0,249). In der Gesamtstichprobe wurde abermals eine nominale Assoziation für das Minor-G-Allel der Variante rs34712024 mit ASS nachgewiesen (OR = 0,409; CI95 = 0,188 – 0,888; p-Wert = 0,018). Auffällig ist, dass die Odds Ratio dieser Variante in allen drei Stichproben bei einem Wert < 1 liegt.

In der Detektionsstichprobe konnte für den trimeren STR **rs71781329** gezeigt werden, dass die Minorallele GCG[7] und GCG[8] mit ASS nominal assoziiert vorliegen (OR = ∞; CI95 = ∞; p-Wert = 0,041). Alle Elternteile, bei denen eines der Minorallele nachgewiesen werden konnte, haben dieses an ihre erkrankten Kinder weitervererbt. Das Ergebnis des TDT konnte weder in der Replikationsstichprobe (OR = 0,486; CI95 = 0,137 – 1,733; p-Wert = 0,239) noch in der Gesamtstichprobe (OR = 0,875; CI95 = 0,317 – 2,417; p-Wert = 0,796) repliziert werden. Die getrennte Berechnung der beiden möglichen Minorallele ergab zwar keine signifikante Assoziation zu ASS, jedoch zeigte sich deskriptiv, dass das Allel rs71781329GCG[7] untertransmittiert (OR = 0,600; CI95 = 0,143 – 2,511; p-Wert = 0,477), rs71781329GCG[8] hingegen übertransmittiert vorliegt (OR = 1,338; CI95 = 0,292 – 6,142; p-Wert = 0,704). Die Ergebnisse der TDTs sind detailliert in Tabelle 9 aufgelistet.

Bei der Post-hoc Poweranalyse zeigte es sich, dass für das Allel rs34712024G die Power der Replikationsstichprobe bei 75% liegt, um die bei der Sequenzierung nachgewiesenen OR von 0,200 replizieren zu können. Für den STR rs71781329 zeigte sich, dass Effekte $0,223 > OR > 4,480$ mit einer Power > 80% nachgewiesen werden können. Bei der Kombination der beiden Stichproben (N = 592) hat sich gezeigt, dass die Power bei > 80% liegt, um für das Allel rs34712024G eine $OR \leq 0,46$ zu detektieren, bzw. um für den STR rs71781329 eine $OR \leq 0,29$ nachzuweisen.

Tabelle 9: Familienbasierte Einzelmarkeranalyse der Stichproben mit UNPHASED.

SNP ^a	GE (%)	MAF ^b	OR	CI 95%	p-Wert ^c
Detektionsstichprobe (N=236 Familien)					
rs150447075T/G	100	0,020	1,000	0,322 – 3,101	1,000
rs34712024A/G	100	0,017	0,200	0,044 – 0,913	0,016*
rs71781329GCG[6]/GCG(7_8)	100	0,005	∞ ^d	∞ ^d	0,041*
Replikationsstichprobe (N=356 Familien)					
rs150447075T/G	99,73	0,025	0,888	0,453 – 1,744	0,732
rs34712024A/G	99,73	0,015	0,583	0,230 – 1,482	0,249
rs71781329GCG[6]/GCG(7_8)	98,37	0,006	0,486	0,137 – 1,733	0,239
Gesamtstichprobe (N=592 Familien)					
rs150447075T/G	99,83	0,023	0,917	0,514 – 1,635	0,768
rs34712024A/G	99,83	0,016	0,409	0,188 – 0,888	0,018*
rs71781329GCG[6]/GCG(7_8)	98,98	0,006	0,875	0,317 – 2,417	0,796
rs71781329GCG[6]/GCG[7] ^e		0,003	0,600	0,143 – 2,511	0,477
rs71781329GCG[6]/GCG[8] ^f		0,003	1,338	0,292 – 6,142	0,704

^aSNP Name mit Major/Minor Allel. ^bMinor-Allelfrequenz innerhalb der Autismusstichprobe. ^cp-Werte < 0,05 sind mit einem Stern markiert. ^dSobald ein Elternteil die Variante trug, wurde diese an das betroffene Kind weitervererbt. ^eAlle Familien, bei denen das Allel rs71781329GCG[8] vorkommt, wurden ausgeschlossen; N = 588 Familien. ^fAlle Familien, bei denen das Allel rs71781329GCG[7] vorkommt, wurden ausgeschlossen; N = 584 Familien. *Abkürzungen:* CI, Konfidenzintervall; GE, Genotypisierungseffizienz in %; OR, Odds Ratio; SNP, Einzelnukleotidpolymorphismus

6.3.2.3 Geschlechtsspezifische und / oder Imprintingeffekte

Es konnten keine signifikanten Imprinting- oder geschlechtsspezifische Effekte nachgewiesen werden (Daten nicht gezeigt). In diesem Zusammenhang ist jedoch zu erwähnen, dass sämtliche Patienten, die das Allel rs34712024G tragen (14 von 591 genotypisierten Proben) männlich sind und dass es nur zwei weibliche Patienten (2 von 585) gibt, bei denen das Minorallel des STR rs71781329 vorliegt.

6.3.2.4 Regressionsanalyse

Um herauszufinden, ob die vorliegenden Genotypen einen Einfluss auf die Sprachentwicklung haben, wurde eine nicht-parametrische Regressionsanalyse durchgeführt. Es ist bekannt, dass

sowohl das Alter zum Zeitpunkt der Diagnose als auch der IQ die Sprachentwicklung beeinflussen, weswegen sie als Kovarianten in das Modell mit eingeflossen sind.

Aufgrund der durchgeführten Regressionsanalyse kann vermutet werden, dass das mit ASS assoziierte Minor-G-Allel der Variante **rs34712024** einen Einfluss (p-Wert = 0,086) auf das Alter hat, in dem die ersten Worte gesprochen wurden (ADI_A9). Wird die gleiche Berechnung mit einer rein männlichen Stichprobe durchgeführt, wird der Effekt etwas deutlicher (p-Wert = 0,067). Für das Alter in dem die ersten Sätze gesprochen wurden (ADI_A10) konnte dies nicht beobachtet werden. Auffällig ist, dass die Patienten, die das bisher protektiv in Erscheinung getretene Allel G tragen, im Vergleich zu den Majorallelträgern (Mittelwert = 25,05 Monate; SD = 14,07; N = 373) rund zehn Monate später anfangen zu sprechen (Mittelwert = 34,80 Monate; SD = 18,50; N = 10). Das Allel scheint sich somit negativ auf das Alter auszuwirken, in dem die ersten Worte gesprochen wurden (Tabelle 10).

Für den STR **rs71781329** zeigte sich, dass das Minorallel GCG[7] sowohl das Alter in dem die ersten Worte gesprochen wurden (ADI_A9; p-Wert < 0,00001) als auch das Alter in dem die ersten Sätze gesprochen wurden (ADI_A10; p-Wert = 0,00001) hochsignifikant beeinflusst, selbst nach Koorektur für multiples Testen (p-Wert (Bonferroni) < 0,0012; Tabelle 10). Es ist jedoch zu beachten, dass diese Variante nur 2mal in der gesamten Stichprobe vorkommt, und somit das Ergebnis mit Vorsicht zu diskutieren ist. Eine nach Geschlechtern getrennte Berechnung wurde aufgrund der geringen Anzahl an Trägern nicht durchgeführt.

Bei den Trägern der Minorallele **rs71781329**GCG[8] und **rs150447075**G ist im Vergleich mit Trägern der entsprechenden Majorallelen kein Unterschied in der Sprachentwicklung nachweisbar (Tabelle 10).

6.4 Funktionelle Analyse der Varianten **rs150447075**, **rs34712024** und **rs71781329**

In einer in-silico Analyse mit der Genomatrix Software MatInspector V2.6 (Quandt et al. 1995) wurde geschaut, ob je nach vorliegendem Allel andere Transkriptionsfaktoren in unmittelbarer Nähe (\pm 50bp) der analysierten Varianten binden. Abschließend erfolgte mittels Luciferase-Assay eine in-

Tabelle 10: Nach Genotypen aufgetrennte Varianten rs150447075, rs34712024 und rs71781329 mit Mittelwerten des Spracherwerbs (ADI_A9 und ADI_A10) und den Ergebnissen der Regressionsanalyse.

		Majorallel (homozygot)		Minorallel		
rs150447075		TT		GT oder GG		
		<i>N</i>	<i>Mittelwert (SD)</i>	<i>N</i>	<i>Mittelwert (SD)</i>	<i>p-Wert</i>
<i>Gemischt</i>	ADI_A9	366	25,48 (14,42)	17	21,65 (9,51)	0,322
	ADI_A10	320	37,69 (17,95)	16	33,94 (10,47)	0,331
<i>Nur männlich</i>	ADI_A9	217	24,82 (14,14)	14	21,93 (10,41)	0,606
	ADI_A10	279	37,01 (18,05)	13	33,77 (10,19)	0,676
rs34712024		AA		AG oder GG		
		<i>N</i>	<i>Mittelwert (SD)</i>	<i>N</i>	<i>Mittelwert (SD)</i>	<i>p-Wert</i>
<i>Gemischt</i>	ADI_A9	373	25,05 (14,07)	10	34,80 (18,50)	0,086
	ADI_A10	328	37,36 (17,52)	8	43,63 (23,80)	0,899
<i>Nur männlich</i>	ADI_A9	321	24,39 (13,76)	10	34,80 (18,50)	0,067
	ADI_A10	284	36,67 (17,60)	8	43,63 (23,80)	0,871
rs71781329		GCG(6_6)		GCG(6_7) oder GCG(6_8)		
		<i>N</i>	<i>Mittelwert (SD)</i>	<i>N</i>	<i>Mittelwert (SD)</i>	<i>p-Wert</i>
<i>Gemischt</i>	ADI_A9	375	25,03 (13,62)	5	45,00 (37,11)	0,101
	ADI_A10	329	37,22 (16,98)	5	58,80 (42,51)	0,446
<i>Nur männlich</i>	ADI_A9	324	24,44 (13,27)	4	44,25 (42,81)	0,272
	ADI_A10	286	36,63 (17,07)	4	55,50 (48,34)	0,843
rs71781329		GCG(6_6)		GCG(6_7)		
		<i>N</i>	<i>Mittelwert (SD)</i>	<i>N</i>	<i>Mittelwert (SD)</i>	<i>p-Wert</i>
<i>Gemischt</i>	ADI_A9	375	25,03 (13,62)	2	78,00 (42,43)	<0,0001
	ADI_A10	329	37,22 (16,98)	2	99,00 (38,18)	0,0001
<i>Nur männlich</i>	ADI_A9	324	24,44 (13,27)	1	108,00 (NA)	^a
	ADI_A10	286	36,63 (17,07)	1	126,00 (NA)	^a
		GCG(6_6)		GCG(6_8)		
		<i>N</i>	<i>Mittelwert (SD)</i>	<i>N</i>	<i>Mittelwert (SD)</i>	<i>p-Wert</i>
<i>Gemischt</i>	ADI_A9	375	25,03 (13,62)	3	^b	^b
	ADI_A10	329	37,22 (16,98)	3	^b	^b
<i>Nur männlich</i>	ADI_A9	324	24,44 (13,27)	3	23,00 (6,25)	0,741
	ADI_A10	286	36,63 (17,07)	3	32,00 (13,86)	0,748

^aBerechnung wurde nicht durchgeführt, da für Minorallel GCG(6_7) nur ein männlicher Patient vorlag. ^bVariante konnte in keinem weiblichen Patienten nachgewiesen werden. *Abkürzungen:* ADI_A9, Alter in dem die ersten Worte gesprochen wurden (in Monaten); ADI_A10, Alter in dem die ersten Sätze gesprochen wurden (in Monaten); N, Anzahl der Allelträger mit vorliegenden ADI-Werten; SD, Standardabweichung.

in vitro Analyse der Varianten rs150447075, rs34712024 und rs71781329 in undifferenzierten HEK293T Zellen und undifferenzierten SH-SY5Y Zellen, sowie 72 und 216 Stunden nach Induktion der neuronalen Differenzierung der SH-SY5Y Zellen mittels Retinsäure und hBDNF (Brain-derived

neurotrophic factor human). Die Auswahl der Messzeitpunkte erfolgt aufgrund des Befundes, dass *CNTNAP2* zu diesen Zeitpunkten maximal reguliert vorliegt. Für Details zur Auswahl der Zeitpunkte siehe Abschnitt 5.5.2.5.3.

6.4.1 In-silico Analyse

Durch das Minor-G-Allel des SNP **rs150447075** entsteht eine zusätzliche Bindestelle für GLI3 (Glioma-associated oncogene family zinc finger 3), einem hirnspezifischen Krüppel like Transkriptionsfaktor, der an der Extremitätenausbildung beteiligt ist. Die Bindestelle für den neuronalen TF NRSF (Neuron-restrictive silencer factor) wird durch das Minorallel nicht verändert.

Durch **rs34712024G** entsteht eine zusätzliche Bindestelle für ZF9, einem Krüppel like Transkriptionsfaktor (TF), welcher dafür bekannt ist, dass er während der Differenzierung und dem Zellwachstum als starker Transkriptionenhancer agiert (DiFeo et al. 2009). Die Bindestellen für den Transkriptionsrepressor ZF300, das bifunktionale Aktivator-/Repressor-Protein TFAP4 und den Transkriptionsaktivator MYOG fallen durch Vorliegen des Minor-G-Allels weg.

Bei Vorliegen der Minorallele GCG[7] und GCG[8] des STR **rs71781329** nimmt die Anzahl pro Repeat um jeweils eine mögliche Bindestelle für den Transkriptionsfaktor EGR1 (Early Growth Response 1) zu. Dieser spielt eine zentrale Rolle in der Regulation der neuronalen Plastizität und der Differenzierung. Die Ergebnisse der in-silico Analyse für die Varianten rs150447075, rs34712024 und rs71781329 sind schematisch in Abbildung 22 dargestellt und in Tabelle 11 detailliert aufgelistet.

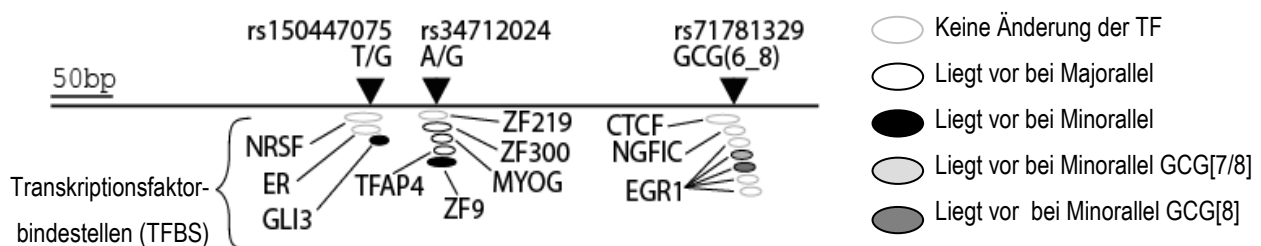


Abbildung 22: Analysierte Varianten, deren schematische Lokalisation innerhalb der *CNTNAP2*-Promotorregion und mögliche Veränderungen der Transkriptionsfaktorbindestellen (TFBS). *Abkürzungen:* TFAP4, Activator protein 4; CTCF, Insulator protein CTCF (CCCTC-binding factor); EGR1, early growth response 1; ER, Estrogen response element; GLI3, Glioma-associated oncogene family zinc finger 3; MYOG, Myogenic bHLH protein myogenin (myf4); NGFIC, Nerve growth factor-induced protein C; NRSF, Neuron-restrictive silencer factor; ZF9, Core promoter-binding protein (CPBP) mit 3 Krüppel-type zinc fingers (auch bekannt als KLF6); ZF219, Krüppel-like zinc finger protein 219; ZF300, KRAB-containing zinc finger protein 300

Tabelle 11: Ergebnis der mit MatInspector durchgeführten Abfrage möglicher Transkriptionsfaktorbindestellen. Die Varianten rs150447075, rs34712024 und rs71781329 sind Bestandteil der Ankersequenz der aufgelisteten Transkriptionsfaktoren.

rs150447075T				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
NRSF	Neuron-restrictive silencer factor	1,000	0,970	acttcAGCAccgcgacagcatcacctccg
ER	Estrogen response elements, IR3 sites	0,809	0,861	caccgcgacagcatCACctccc
rs150447075G				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
NRSF	Neuron-restrictive silencer factor	1,000	0,970	acttcAGCAccgcgacagcatcaccgccc
ER	Estrogen response elements, IR3 sites	0,809	0,835	caccgcgacagcatCACGccc
GLI3 ^b	GLI-Kruppel family member GLI3	0,801	0,880	ccgcCCGccaggga
rs34712024A				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
ZF219	Kruppel-like zinc finger protein 219	1,000	0,996	ccttgctccgCCCCgccccagc
ZF300 ^c	KRAB-containing zinc finger protein 300	1,000	1,000	tgctccgccccgCCCCagctga
MYOG ^c	Myogenic bHLH protein myogenin (myf4)	1,000	0,908	ccgcccCAGCtgatgga
TFAP4 ^c	Activator protein 4	1,000	0,949	cgccccAGCTgatggac
rs34712024G				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
ZF219	Kruppel-like zinc finger protein 219	1,000	0,996	ccttgctccgCCCCgccccggc
ZF9 ^b	Core promoter-binding protein (CPBP) with 3 Kruppel-type zinc fingers (KLF6)	1,000	0,925	ttgctccgcccCGCccccggctg
rs71781329GCG[6]				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
CTCF	Insulator protein CTCF (CCCTC-binding factor)	0,863	0,796	ccgggctcacaggagcTG GCggcgggcg
NGFIC	Nerve growth factor-induced protein C	0,786	0,812	gctg GCGGcgggcgcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcgt
EGR1	Early growth response 1	0,789	0,894	ggcgggcgCGGCgtctc
rs71781329GCG[7]				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
CTCF	Insulator protein CTCF (CCCTC-binding factor)	0,863	0,796	ccgggctcacaggagcTG GCggcgggcg
NGFIC	Nerve growth factor-induced protein C	0,786	0,812	gctg GCGGcgggcgcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcgt
EGR1 ^b	Early growth response 1	0,789	0,894	ggcgggcgCGGCgtctc
rs71781329GCG[8]				
TF	Vollständiger Name	Kern sim.	Matrix sim.	Ankersequenz ^a
CTCF	Insulator protein CTCF (CCCTC-binding factor)	0,863	0,796	ccgggctcacaggagcTG GCggcgggcg
NGFIC	Nerve growth factor-induced protein C	0,786	0,812	gctg GCGGcgggcgcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcg
EGR1 ^b	Early growth response 1	0,789	0,910	ggcgggcgCGGCggcgt
EGR1 ^b	Early growth response 1	0,789	0,894	ggcgggcgCGGCgtctc

^aMatrix der Transkriptionsfaktorbindestellen; die Ankersequenz ist in Großbuchstaben dargestellt; die Varianten sind im Text fett hervorgehoben. ^bHinzugekommene Transkriptionsfaktorbindestelle (kursiv). ^cWeggefallene Transkriptionsfaktorbindestelle (kursiv). **Abkürzungen:** TFAP4, Activator protein 4; CTCF, Insulator protein CTCF (CCCTC-binding factor); EGR1, early growth response 1; ER, Estrogen response element; GLI3, Glioma-associated oncogene family zinc finger 3; MYOG, Myogenic bHLH protein myogenin (myf4); NGFIC, Nerve growth factor-induced protein C; NRSF, Neuron-restrictive silencer factor; sim, Ähnlichkeit; TF, Transkriptionsfaktor; ZF9, Core promoter-binding protein (CPBP) mit 3 Kruppel-type Zinkfinger (auch bekannt als KLF6); ZF219, Kruppel-like zinc finger protein 219; ZF300, KRAB-containing zinc finger protein 300.

6.4.2 In-vitro Analyse

6.4.2.1 Transfektionseffizienz-Test

Um die best mögliche Transkriptionseffizienz für die Neuroblastomzelllinie SH-SY5Y herauszufinden, wurden Transfektionsreagenzien mehrerer Hersteller unter verschiedenen Bedingungen an der Zelllinie ausgetestet.

Die höchste Transfektionseffizienz von 29,9% wurde von dem Reagenz Metafectene® Pro (Biontex Laboratories GmbH) erreicht. Bei der weiteren Optimierung der Transfektionsbedingungen zeigte sich, dass bei diesem Reagenz der Reaktionsansatz mit PBS zu einer weiteren Erhöhung der Effizienz führt. Die Transfektion der HEK293T Zellen sowie sämtliche Transfektionsreaktionen der SH-SY5Y Zellen wurden daher mit diesem Reagenz durchgeführt. Für den exakten Reaktionsansatz siehe Abschnitt 5.5.2.5.4.2.

6.4.2.1.1 Attractene (Qiagen)

Die ermittelte Transfektionseffizienz für das Reagenz Attractene (Qiagen) lag je nach DNA-Konzentration und Verhältnis von DNA zu Reagenz zwischen 2,5 – 16,7% (Tabelle 12). Die maximale Transfektionseffizienz von 16,7% wurde bei einem Verhältnis (DNA:Reagenz) von 1:7,5 erreicht.

Tabelle 12: Transfektionseffizienztestergebnis für das Reagenz Attractene.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,2µg / 0,5µl	1:1,5	283	7	2,5
0,4µg / 1,0µl	1:1,5	252	12	4,8
0,6µg / 1,5µl	1:1,5	227	12	5,3
0,2µg / 0,75µl	1:3,75	243	14	5,8
0,4µg / 1,5µl	1:3,75	150	17	11,3
0,6µg / 2,25µl	1:3,75	129	10	7,8
0,2µg / 1,5µl	1:7,5	167	10	6,0
0,4µg / 3,0µl	1:7,5	114	14	12,3
0,6µg / 4,5µl	1:7,5	66	11	16,7

6.4.2.1.2 Metafectene® Easy+ (Biontex Laboratories GmbH)

Die ermittelte Transfektionseffizienz für das Reagenz Metafectene® Easy+ (Biontex Laboratories GmbH) lag je nach eingesetztem Volumen des DNA-Transfektionsreagenz-Gemischs zwischen 15,0 – 20,2% (Tabelle 13). Die maximale Transfektionseffizienz von 20,2% wurde bei einem Einsatz von

1,3µg DNA und 1,3µl Transfektionsreagenz, bzw. 1,9µg DNA und 1,9µl Transfektionsreagenz erreicht.

Tabelle 13: Transfektionseffizienztestergebnis für das Reagenz Metafectene® Easy*.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,6µg / 0,6µl	1:1	453	68	15,0
1,3µg / 1,3µl	1:1	262	53	20,2
1,9µg / 1,9µl	1:1	213	43	20,2

6.4.2.1.3 Metafectene® Pro (Biontex Laboratories GmbH)

Die ermittelte Transfektionseffizienz für das Reagenz Metafectene® Pro (Biontex Laboratories GmbH) lag je nach DNA-Konzentration und Verhältnis von DNA zu Reagenz zwischen 3,0 – 29,9% (Tabelle 14). Die maximale Transfektionseffizienz von 29,9% wurde bei einem Verhältnis (DNA:Reagenz) von 1:4 erreicht.

Tabelle 14: Transfektionseffizienztestergebnis für das Reagenz Metafectene® Pro.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,5µg / 0,5µl	1:1	168	8	4,8
1,0µg / 1,0µl	1:1	189	18	9,5
1,5µg / 1,5µl	1:1	474	14	3,0
0,5µg / 1,0µl	1:2	229	17	7,4
1,0µg / 2,0µl	1:2	176	27	15,3
1,5µg / 3,0µl	1:2	479	40	8,4
0,5µg / 2,0µl	1:4	315	67	21,3
1,0µg / 4,0µl	1:4	231	69	29,9
1,5µg / 6,0µl	1:4	234	55	23,5
0,5µg / 3,0µl	1:6	146	14	9,6
1,0µg / 6,0µl	1:6	139	21	15,1
1,5µg / 9,0µl	1:6	120	24	20,0

6.4.2.1.4 Fugene® HD (Promega GmbH)

Die ermittelte Transfektionseffizienz für das Reagenz Fugene® HD (Promega GmbH) lag je nach DNA-Konzentration und Verhältnis von DNA zu Reagenz zwischen 0 – 8,6% (Tabelle 15). Die maximale Transfektionseffizienz von 8,6% wurde bei einem Verhältnis (DNA:Reagenz) von 1:2 erreicht. Die Ansätze mit einer Transfektionseffizienz von 0% wurden nicht ausgewertet.

Tabelle 15: Transfektionseffizienztestergebnis für das Reagenz Eugene® HD.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,5µg / 0,5µl	1:1	NA	0	0
1,0µg / 1,0µl	1:1	NA	0	0
1,5µg / 1,5µl	1:1	NA	0	0
0,5µg / 1,0µl	1:2	275	7	2,5
1,0µg / 2,0µl	1:2	514	31	6,0
1,5µg / 3,0µl	1:2	256	22	8,6
0,5µg / 1,5µl	1:3	232	6	2,6
1,0µg / 3,0µl	1:3	364	16	4,4
1,5µg / 4,5µl	1:3	550	15	2,7
0,5µg / 2,0µl	1:4	326	5	1,5
1,0µg / 4,0µl	1:4	189	2	1,1
1,5µg / 6,0µl	1:4	605	5	0,8

6.4.2.1.5 Xfect™ (Takara Bio Europe SAS)

Die ermittelte Transfektionseffizienz für das Reagenz Xfect™ (Takara Bio Europe SAS) lag je nach DNA-Konzentration und eingesetzter Menge des Reagenzes zwischen 11,9 – 25,0% (Tabelle 16). In den zuvor dicht besiedelten Regionen sind bei diesem Reagenz sehr viele Zellen gestorben. Die Transfektionseffizienzrate wurde daher ausschließlich in den dünner besiedelten Randgebieten des Wells bestimmt. Die maximale Transfektionseffizienz von 25,0% wurde bei einem Verhältnis (DNA:Reagenz) von 1:2 erreicht.

Tabelle 16: Transfektionseffizienztestergebnis für das Reagenz Xfect™.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,5µg / 0,15µl	1:0,3	84	10	11,9
1,0µg / 0,3µl	1:0,3	178	28	15,7
1,5µg / 0,45µl	1:0,3	84	21	25,0

6.4.2.1.6 TransIT®-2020 (Mirus Bio LLC)

Die ermittelte Transfektionseffizienz für das Reagenz TransIT®-2020 (Mirus Bio LLC) lag je nach DNA-Konzentration und Verhältnis von DNA zu Reagenz zwischen 1,3 – 12,3% (Tabelle 17). Die maximale Transfektionseffizienz von 12,3% wurde bei einem Verhältnis (DNA:Reagenz) von 1:4 erreicht.

Tabelle 17: Transfektionseffizienztestergebnis für das Reagenz TransIT®-2020 (Mirus Bio LLC).

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,5µg / 0,5µl	1:1	594	10	1,7
1,0µg / 1,0µl	1:1	332	9	2,7
1,5µg / 1,5µl	1:1	546	8	1,5
0,5µg / 1,0µl	1:2	311	4	1,3
1,0µg / 2,0µl	1:2	365	5	1,4
1,5µg / 3,0µl	1:2	484	16	3,3
0,5µg / 1,5µl	1:3	599	14	2,3
1,0µg / 3,0µl	1:3	303	24	7,9
1,5µg / 4,5µl	1:3	430	30	7,0
0,5µg / 2,0µl	1:4	379	19	5,0
1,0µg / 4,0µl	1:4	539	18	3,3
1,5µg / 6,0µl	1:4	285	35	12,3

6.4.2.1.7 TurboFect (Fermentas GmbH)

Die ermittelte Transfektionseffizienz für das Reagenz TurboFect (Fermentas GmbH) lag je nach DNA-Konzentration und Verhältnis von DNA zu Reagenz zwischen 3,2 – 20,6% (Tabelle 18). Die maximale Transfektionseffizienz von 20,6% wurde bei einem Verhältnis (DNA:Reagenz) von 1:3 erreicht.

Tabelle 18: Transfektionseffizienztestergebnis für das Reagenz TurboFect.

DNA [µg] / Reagenz [µl]	Verhältnis DNA:Reagenz	Zellen gesamt	Zellen transfiziert	Prozent
0,5µg / 0,5µl	1:1	215	13	6,0
1,0µg / 1,0µl	1:1	318	27	8,5
1,5µg / 1,5µl	1:1	215	10	4,7
0,5µg / 1,0µl	1:2	422	21	5,0
1,0µg / 2,0µl	1:2	701	35	5,0
1,5µg / 3,0µl	1:2	656	21	3,2
0,5µg / 1,5µl	1:3	254	33	13,0
1,0µg / 3,0µl	1:3	238	31	13,0
1,5µg / 4,5µl	1:3	399	82	20,6
0,5µg / 2,0µl	1:4	306	57	18,6
1,0µg / 4,0µl	1:4	684	56	8,2
1,5µg / 6,0µl	1:4	215	13	6,0

6.4.2.2 Luciferase-Assay

Um zu testen, ob die drei Promotorvarianten die Transkriptionseffizienz beeinflussen, wurde ein Luciferase-Assay durchgeführt. Mit dieser Methode ist es möglich, die Auswirkung von Promotorvarianten auf die Genregulation zu untersuchen.

Die ersten Vorversuche in der fibroblastoiden Zelllinie HEK293T wurden in Kooperation mit Frau Dipl. Biol. Ulrike Graab am Institut für Experimentelle Tumorforschung in der Pädiatrie (Frankfurt/Main) durchgeführt (Daten nicht gezeigt). Auf Grund dieser ersten erfolgreichen Experimente wurde die Methode am Molekulargenetischen Labor der Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes- und Jugendalters (Frankfurt/Main) etabliert und weiter optimiert. Die Ergebnisse der ersten Messungen wurden erfolgreich repliziert und die Methode auf die neuronale Zelllinie SH-SY5Y ausgedehnt. Hier werden nur eigene Ergebnisse aufgeführt und diskutiert.

6.4.2.2.1 Luciferaseaktivität in der Zelllinie HEK293T

Ein nominal signifikanter Anstieg der Promotoraktivität konnte sowohl für das G-Allel der Variante **rs150447075** (1,07fach; SD = 0,07; rANOVA p-Wert = 0,050), als auch für das mit ASS assoziierte G-Allel der Variante **rs34712024** (1,11fach; SD = 0,07; rANOVA p-Wert = 0,038) im Vergleich zum Promotor mit den Majorallelkonstrukt nachgewiesen werden. Für die Minorallele der Variante **rs71781329** konnte gezeigt werden, dass die Transkriptionseffizienz des Promotors je nach Anzahl der Tripletts ansteigt. Im Vergleich zum Majorallel GCG[6] zeigt das Minorallel GCG[7] keinen Anstieg der Promotoraktivität (1,01fach; SD = 0,02; rANOVA p-Wert = 0,883), wohingegen die Aktivität des Promotors durch das Vorliegen des Minorallel GCG[8] nominal signifikant erhöht ist (1,19fach; SD = 0,03; rANOVA p-Wert = 0,031). Die Ergebnisse des Luciferaseassay sind detailliert in Tabelle 19 aufgelistet.

6.4.2.2.2 Luciferaseaktivität in der Zelllinie SH-SY5Y

Bei der SH-SY5Y Zelllinie war zu beobachten, dass die Signalstärke bei der Luciferaseaktivitätsmessung wesentlich geringer ausfiel, was auf die reduzierte Transfektionseffizienz bei dieser Zelllinie ($\leq 50\%$) zurückzuführen ist. Gemessen wurde die Luciferaseaktivität der SH-SY5Y Zelllinie zu den Zeitpunkten 1 (Undifferenziert), 3 (72 Stunden in Differenzierungsmedium - Differenziert) und 6 (216 Stunden in Differenzierungsmedium - Ausdifferenziert). Die Auswahl der Zeitpunkte erfolgte aufgrund einer mRNA-Expressionsstudie von Frau Denise Haslinger (Msc). Hierbei konnte Sie zeigen, dass die Expression von *CNTNAP2* in den ersten Tagen der Differenzierung runterreguliert wird und zum Zeitpunkt 3 seinen tiefsten Punkt erreicht. Zum Zeitpunkt 6, also zum Ende der Zelldifferenzierung normalisiert sich die *CNTNAP2*-Expressionsrate

wieder. Für detailliertere Angaben zur Auswahl der Messzeitpunkte siehe Abschnitt 5.5.2.5.3. Die Ergebnisse des Luciferaseassays sind detailliert in Tabelle 19 aufgelistet.

Tabelle 19: Relative Luciferaseaktivität des *CNTNAP2*-Promotors in den Zelllinien HEK293T und SH-SY5Y.

HEK293T	Mittelwert ^a	SD (Mittelwert)	Verhältnis ^b	SD (Verhältnis)	p-Wert
Leervektor	1	0,05			
Majorallelkonstrukt	7,12	0,54	1,00	0,08	
rs150447075G	7,63	0,48	1,07	0,07	0,050*
rs34712024G	7,93	0,53	1,11	0,07	0,038*
rs71781329GCG[7]	7,15	0,12	1,01	0,02	0,883
rs71781329GCG[8]	8,44	0,20	1,19	0,03	0,031*
SH-SY5Y - 0h	Mittelwert ^a	SD (Mittelwert)	Verhältnis ^b	SD (Verhältnis)	p-Wert
Leervektor	1	0,02			
Majorallelkonstrukt	5,20	0,13	1,00	0,03	
rs150447075G	5,57	0,17	1,07	0,03	0,008**
rs34712024G	5,05	0,22	0,97	0,04	0,105
rs71781329GCG[7]	4,89	0,20	0,94	0,04	0,058
rs71781329GCG[8]	5,11	0,04	0,98	0,01	0,371
SH-SY5Y - 72h	Mittelwert ^a	SD (Mittelwert)	Verhältnis ^b	SD (Verhältnis)	p-Wert
Leervektor	1	0,07			
Majorallelkonstrukt	6,27	0,40	1,00	0,06	
rs150447075G	6,95	0,42	1,11	0,07	0,234
rs34712024G	5,75	0,48	0,92	0,08	0,443
rs71781329GCG[7]	6,49	0,21	1,04	0,03	0,539
rs71781329GCG[8]	5,69	0,23	0,91	0,04	0,022*
SH-SY5Y - 216h	Mittelwert ^a	SD (Mittelwert)	Verhältnis ^b	SD (Verhältnis)	p-Wert
Leervektor	1	0,34			
Majorallelkonstrukt	45,60	20,47	1,00	0,45	
rs150447075G	52,59	25,14	1,18	0,55	0,455
rs34712024G	45,90	17,06	1,05	0,37	0,808
rs71781329GCG[7]	49,88	23,34	1,12	0,51	0,565
rs71781329GCG[8]	42,03	19,14	0,93	0,42	0,623

^aMittelwert der Luciferaseaktivität der drei zum Leervektor normalisierten biologischen Replikate. ^bVerhältnis der relativen Luciferaseaktivität des Majorallelkonstrukts und der entsprechenden Minorallele. Sterne weisen auf signifikante Unterschiede in der Promotoraktivität zwischen den Majorallelen und den jeweiligen Minorallelen hin (* p-Wert < 0,050; ** p-Wert < 0,010).

Abkürzungen: h, Stunden; SD, Standardabweichung.

6.4.2.2.1 Zeitpunkt 1 – Undifferenziert

Für das nicht mit ASS assoziierte G-Allel der Variante **rs150447075** konnte ein 1,07facher (SD = 0,03), signifikanter Anstieg der Promotoraktivität nachgewiesen werden (rANOVA p-Wert = 0,008). Hingegen konnte weder für das mit ASS assoziierte Allel **rs34712024G**, noch für die Allele der Variante **rs71781329** eine signifikante Änderung nachgewiesen werden. Im Vergleich zum Minorallel

GCG[8] (0,98fach; SD = 0,04; rANOVA p-Wert = 0,371) zeigt das Minorallel GCG[7] jedoch einen Trend in der Regulierung der Promotoraktivität (0,94fach; SD = 0,01; rANOVA p-Wert = 0,058).

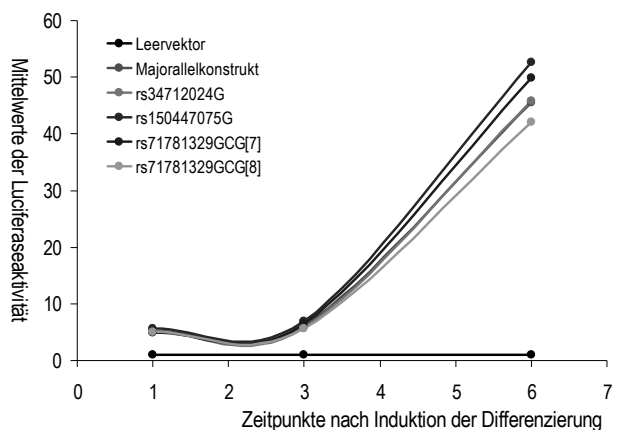
6.4.2.2.2 Nach 72 Stunden – Während der Differenzierung

Es konnte keine signifikante Änderung der Promotoraktivität für die G-Allele der Varianten **rs150447075** (1,11fach; SD = 0,07; rANOVA p-Wert = 0,234) und **rs34712024** (0,92fach; SD = 0,08; rANOVA p-Wert = 0,443) und das Minorallel GCG[7] der Variante **rs71781329** (1,04fach; SD = 0,03; rANOVA p-Wert = 0,539) nachgewiesen werden. Das Minorallel GCG[8] hingegen zeigt im Vergleich zum Majorallel GCG[6] eine signifikante Runterregulierung der Promotoraktivität (0,91fach; SD = 0,04; rANOVA p-Wert = 0,022).

6.4.2.2.3 Nach 216 Stunden – Ausdifferenziert

Für keines der Allele rs150447075G, rs34712024G, rs71781329GCG[7] und rs71781329GCG[8] konnte eine signifikante Änderung der Promotoraktivität nachgewiesen werden (rANOVA p-Wert \geq 0,455). Zu diesem Zeitpunkt sind die gegen den Leervektor normalisierten Mittelwerte der Messung und die Standardabweichung im Vergleich zu den Messzeitpunkten 0h und 72h deutlich erhöht (siehe Abbildung 23 und Tabelle 19).

Abbildung 23: Normalisierte Mittelwerte (Referenzvektor = Leervektor) der Luciferaseaktivität der analysierten Varianten zu unterschiedlichen Messzeitpunkten.



6.5 Zusammenfassung der Ergebnisse

In der Detektionsstichprobe (N = 236) konnten durch direkte Sequenzierung der *CNTNAP2*-Promotorregion die Varianten rs150447075, rs34712024 und rs71781329 nachgewiesen werden. Zusätzlich wurden sechs bisher nicht beschriebene Varianten entdeckt. Die drei häufigsten Varianten (Genotypfrequenz > 1%) wurden zur Erweiterung der Stichprobe in weiteren 356 Familien (Replikationsstichprobe) genotypisiert.

Bei der statistischen Auswertung der Genotypisierungsdaten hat sich gezeigt, dass das Minor-G-Allel der Variante rs34712024 sowohl in der Detektionsstichprobe, als auch in der Gesamtstichprobe nominal mit ASS assoziiert vorliegt. In der Replikationsstichprobe konnte dieses Ergebnis, möglicherweise aufgrund der zu geringen Power, nicht repliziert werden. In allen drei Stichproben liegt die Odds-Ratio dieser Variante bei einem Wert < 1 .

Für die Minorallele GCG[7] und GCG[8] des trimeren STR rs71781329 konnte ebenfalls eine nominale Assoziation mit ASS in der Detektionsstichprobe gezeigt werden, die jedoch weder in der Replikations- noch in der Gesamtstichprobe repliziert werden konnte. Die getrennte Berechnung der beiden Minorallele zeigte, dass das Allel rs71781329GCG[7] untertransmittiert (OR = 0,600), rs71781329GCG[8] hingegen übertransmittiert in der Gesamtstichprobe vorliegt (OR = 1,338), wobei hier ein Zufallsbefund nicht ausgeschlossen werden kann.

Bei der funktionellen in-silico Analyse der Varianten rs150447075, rs34712024 und rs71781329 zeigte sich, dass sich je nach vorliegendem Allel die Sequenz so ändert, dass in unmittelbarer Nähe (± 50 bp) bindende Transkriptionsfaktoren wegfallen bzw. neue Bindestellen entstehen. Bei der abschließenden in-vitro Analyse der Varianten mittels Luciferase-Assay konnte sowohl ein zelllinienspezifischer Effekt, als auch ein vom Entwicklungsstatus der Neuroblastomzelllinie abhängiger Effekt der Varianten nachgewiesen werden.

7 Diskussion

Im Rahmen dieser Studie wurde gezeigt, dass die häufigsten Varianten rs150447075, rs34712024 und rs71781329 der Promotorregion des ASS Risikogens *CNTNAP2* funktionelle Effekte auf die Promotoraktivität haben und rs34712024 sowie rs71781329 mit dem Risiko für ASS bzw. der Sprachentwicklung in ASS assoziiert sind.

7.1 Analyisierte Stichprobe

In unserer analysierten Stichprobe waren 86% der Patienten männlich und 14% weiblich, was einem Geschlechterverhältnis von 6,1:1 (männlich:weiblich) entspricht. Dies ist zwar in etwa vergleichbar mit dem in der Literatur beschriebenen Geschlechterverhältnis von durchschnittlich 3-4:1 (Volkmar et al. 2004), dennoch ist das Verhältnis der vorliegenden Stichprobe leicht zugunsten der männlichen Probanden verschoben. Eine solche Verschiebung ist in Stichproben zu beobachten, die

überwiegend Probanden mit einer milderen Form von ASS enthalten, wobei sich hierbei das männlich:weiblich Verhältnis zugunsten der männlichen Probanden, bis zu 11:1 (Gillberg et al. 2006) und mehr, verschieben kann. In Stichproben die sich überwiegend aus schweren Fällen der ASS zusammensetzen, kann dieses Verhältnis hingegen auf 2:1 sinken (Fombonne 2002). Dies ist darin begründet, dass Mädchen mit ASS meist auch eine schwerere geistige Behinderung aufweisen (Lord und Schopler 1985). Aus epidemiologischen Studien geht hervor, dass die Anzahl der Kinder mit einer milden Form der ASS (HFA + Spektrumsstörung) im Allgemeinen 1,75fach höher ist, als die Anzahl der Kinder mit einer strikten Autismusdiagnose (LFA; Fombonne 2002). In der hier analysierten Stichprobe wurde bei 87,5% der Patienten eine strikte Autismusdiagnose gestellt, welche sich zu 61% aus High- (HFA) und zu 39% aus Low-Functioning Autisten (LFA) zusammensetzt. Fasst man die Patienten mit HFA mit den übrigen 12,5% der Patienten mit Asperger-Syndrom oder atypischem Autismus (PDD-NOS) in die Gruppe der milderen Formen von ASS zusammen, so ergibt sich in unserer Stichprobe daraus ein Verhältnis von milderen Formen von ASS zu LFA von 1,9:1. Auch dieser Wert deutet darauf hin, dass die analysierte Stichprobe mehr Patienten mit milderen Formen des Autismusspektrums beinhaltet und bestätigt somit die in der Stichprobe vorliegende leichte Verschiebung des Geschlechterverhältnisses zugunsten der männlichen Patienten. Ein Grund dafür könnte darin liegen, dass Patienten mit einem IQ < 35 bereits im Voraus von der Studie ausgeschlossen wurden, da in diesen Fällen eine Diagnosestellung mit dem Diagnostischen Interview für Autismus - Revidiert (ADI-R; Lord et al. 1994) bzw. der Beobachtungsskala für Autistische Störungen (ADOS-G; Lord et al. 2000) nicht möglich ist.

7.2 Analyse der Detektionsstichprobe: Sequenzierung der *CNTNAP2*-Promotorregion

Bei der Sequenzierung der *CNTNAP2*-Promotorregion konnten in den 236 Familien der Detektionsstichprobe nur die Varianten rs150447075, rs34712024 und rs71781329 nachgewiesen werden. Die anderen bereits beschriebenen Varianten rs13225855, rs13225866, rs13225872, rs13226652, rs13225963, rs13225979 und rs13226843 wurden in keiner der Proben detektiert. Interessanterweise liegen auch in den frei zugänglichen Datenbanken (dbSNP, HapMap und UCSC; siehe Abschnitt 10.2) nur die Verteilungsdaten für SNP rs150447075 und rs34712024 vor. Die übrigen Varianten wurden somit zwar bereits detektiert, aber nicht im Detail untersucht, obwohl die 5'-liegende Promotorregion eines Gens als primäres Element der Genregulation gilt und Varianten

innerhalb dieser Region die Transkriptionseffizienz des jeweiligen Gens signifikant beeinflussen können (Lesch et al. 1996; Überblick in: Shastri 2009). Durch die Sequenzierung konnten zudem sechs bisher nicht beschriebene Varianten nachgewiesen werden. Da deren Genotypfrequenz in der Detektionsstichprobe aber unter 1% lag und alle von den Eltern an das betroffene Kind vererbt wurden, sind sie in den anschließenden Analysen nicht weiter untersucht worden. Es ist zwar nicht auszuschließen, dass die Varianten einen Effekt auf die Promotoraktivität oder das Risiko für ASS haben, allerdings war aufgrund der geringen Power, bedingt durch das seltene Vorkommen, die Durchführung einer Assoziationsstudie nicht zielführend.

Für den SNP rs150447075 konnte in den erkrankten Kindern unserer Detektionsstichprobe eine Minorallelfrequenz (MAF) von 0,023 berechnet werden, was mit der Allelfrequenz der Gesamtbevölkerung von 0,022 übereinstimmt (www.ncbi.nlm.nih.gov/sites/SNP). Das Minorallel der Variante rs34712024 kam allerdings etwas seltener vor (0,011; CI95 = 0,002 - 0,020) als in den Datenbanken angegeben (0,029; CI95 = 0,006 - 0,052; www.ncbi.nlm.nih.gov/sites/SNP). Diese Abweichung ist grundsätzlich annehmbar, da diese Variante bei weniger als 10% der Personen in der Gesamtbevölkerung vorkommt und sich die CI95 der MAF der beiden Stichproben deutlich überschneiden. Auch wenn bei Trio basierten Assoziationsstudien Populationsstratifikationseffekte keine Rolle spielen ist es dennoch nicht auszuschließen, dass möglicherweise eine Populationsstratifikation zwischen den veröffentlichten Daten und der hier untersuchten Stichprobe vorliegt. Daher sollten bei einer Fall-Kontroll-Studie dieser Variante eventuelle populationsbedingte Effekte berücksichtigt werden. Nur so kann festgestellt werden, ob die von uns beobachtete Abweichung der MAF den normalen Schwankungen unterliegt, oder ob sie spezifisch auf die ASS Stichprobe zurückzuführen ist.

Besonders spannend ist, dass für den STR rs71781329 nicht nur der bereits beschriebene (www.ncbi.nlm.nih.gov/sites/SNP) 8fache Repeat (GCG[8]) dieser Variante nachgewiesen werden konnte, sondern zusätzlich eine 7fache Wiederholung (GCG[7]) des GCG-Mikrosatelliten. Mikrosatelliten können aufgrund ihres gehäuftten Vorkommens in Promotorregionen an der Regulation der Genexpression beteiligt sein und können so den vorliegenden Phänotypen beeinflussen (Sawaya et al. 2013). Diese Tatsache und die unterschiedliche (deskriptive) Transmission der beiden beobachteten Allele des STR rs71781329 in der Gesamtstichprobe lassen daher vermuten, dass sie die autismusspezifischen Phänotypen in unterschiedlicher Weise

beeinflussen. Eine ähnliche Beobachtung konnte auch im funktionellen Ansatz der vorliegenden Arbeit gemacht werden (siehe Abschnitt 7.5).

7.3 Analyse der Replikationsstichprobe: Genotypisierung und statistische Auswertung der Daten

Im Transmissions Disequilibrium Test (TDT) konnte nachgewiesen werden, dass für das Minorallel der Varianten rs34712024 eine Assoziation mit ASS vorliegt und dass das Allels rs71781329GCG[7] die Sprachentwicklung negativ beeinflusst.

Für die Variante rs150447075 konnte in keiner der Stichproben eine Assoziation mit ASS nachgewiesen werden. Auch nach der Regressionsanalyse liegt für diese Variante kein signifikantes Ergebnis vor. Die Variante scheint somit weder das Risiko für ASS noch den Sprachphänotypen in ASS zu beeinflussen. Zudem kam diese Variante in der Stichprobe mit einer MAF > 2 % vor, sodass die Power für diese Stichprobe ausreichend war (Power > 80%), um signifikante (p-value < 0,05) Ergebnisse bei einer minimalen OR von 0,490 in der gesamten Stichprobe zu erhalten. Ein falsch negativer Befund kann somit entweder ausgeschlossen werden, oder der Effekt der Variante ist wesentlich geringer als 0,490.

Sowohl in der Detektionsstichprobe, als auch in der Gesamtstichprobe zeigte sich, dass das Minor-G-Allel der Variante rs37412024 mit ASS signifikant assoziiert vorliegt (p-Wert $\leq 0,018$) und zudem weniger häufig von den Eltern an das betroffene Kind weitergegeben wurde (OR < 0,409). Es scheint daher einen protektiven Effekt auszuüben. Die Tatsache, dass es in der Gesamtstichprobe mit einer Allelfrequenz von 0,016 nachgewiesen werden konnte und somit in der analysierten Stichprobe seltener vorkommt als in den Datenbanken angegeben (0,029; www.ncbi.nlm.nih.gov/sites/SNP) könnte ebenso auf die protektive Rolle dieses Alleles hinweisen.

Bei der parameterfreien Regressionsanalyse der Daten bezüglich des Spracherwerbs der Kinder, welche im ADI-R Elterninterview retrospektiv abgefragt werden, hat sich zudem gezeigt, dass die Minorallelträger (rs37412024G) vermutlich später anfangen zu sprechen als die homozygoten Majorallelträger. Dies könnte bedeuten, dass das Allel rs37412024G zusätzlich modulierend auf die Sprachentwicklung einwirkt. Bereits in früheren Studien konnte mehrfach eine Assoziation von intronischen *CNTNAP2*-Varianten mit dem Sprachphänotypen nachgewiesen werden. So zeigten

Alarcón und Kollegen (2008) in ihrer Studie mit einer ASS Stichprobe, dass das Major-C-Allel der intronischen Variante rs2710102 mit dem Sprachphänotypen „Alter in dem die ersten Worte gesprochen wurden“ assoziiert vorliegt und diesen negativ beeinflusst. Auch in einer Untersuchungen an einer australischen Stichprobe („Raine study“) konnte gezeigt werden, dass *CNTNAP2*-Varianten einen Einfluss auf den Sprachphänotypen haben und somit an der Entstehung von spezifischen Sprachentwicklungsstörungen (SSES) beteiligt sein könnten (Whitehouse et al. 2011). Der modulierende Einfluss des Minor-G-Allels der Variante rs37412024 auf die Sprachentwicklung wird zudem durch die Ergebnisse einer früheren Genexpressionsstudie bestärkt. Hierbei konnte gezeigt werden, dass sich während der Hirnentwicklung die mRNA von *CNTNAP2* in Regionen anreichert, die an der Sprachentwicklung beteiligt sind (Alarcón et al. 2008). Spannend ist in diesem Zusammenhang, dass der von uns nachgewiesene Effekt auf den Sprachphänotypen in der rein männlichen Stichprobe trotz der geringeren Power deutlicher wird. Ein geschlechts-spezifischer Effekt, wie er für verschiedene Varianten auch bereits in anderen Studien nachgewiesen werden konnte (Alarcón et al. 2008; Arking et al. 2008) ist daher nicht auszuschließen.

Der negative Effekt des Minor-G-Allels der Variante rs37412024 auf den Sprachphänotypen steht jedoch mit seiner protektiven Rolle bezüglich des ASS Risikos im starken Widerspruch. Interessanterweise zeigte sich in früheren Studien, dass sowohl das Major- als auch das Minorallel einer anderen *CNTNAP2*-Variante (rs2710102), abhängig von der untersuchten Störung als Risikoallele für Sprachentwicklung fungieren. So schließen beispielsweise Kinder mit spezifischer Sprachentwicklungsstörung (SSES), die das Major-C-Allel des SNPs rs2710102 tragen, in einem Test, bei dem sie Phantasieworte mit 1-5 Silben wiederholen müssen (Non-word Repetition Test) deutlich schlechter ab, als Träger des Minor-T-Allels (Vernes et al. 2008). Im Gegensatz hierzu konnte in einer Stichprobe mit Dyslexiepatienten gezeigt werden, dass Träger des Minor-T-Allel im „Non-word Repetition Test“ schlechter abschneiden als die Major-C-Allelträger (Peter et al. 2011; Überblick in: Peñagarikano und Geschwind 2012). Dies deutet darauf hin, dass diese Allele ebenso bidirektional in Abhängigkeit von den analysierten Phänotypen sind. Es kann daher vermutet werden, dass das Minor-G-Allel der hier untersuchten Variante rs37412024 auf der einen Seite vor der Entstehung von ASS schützt, gleichzeitig jedoch einen negativen Einfluss auf den Sprachphänotypen ausübt. Des Weiteren wäre es auch möglich, dass in einem anderen Krankheitsbild nicht das G-Allel sondern das Major-A-Allel dieser Variante verschlechternd auf den Sprachphänotypen einwirkt. Um den Effekt der Allele auf den Sprachphänotypen und dessen

Abhängigkeit vom genetischen Hintergrund und möglicher personenspezifischer Umweltfaktoren genauer zu bestimmen, ist eine weitere Genotypisierung der Variante in unabhängigen Stichproben, inklusive gesunder Kontrollgruppen nötig.

Für den Short Tandem Repeat (STR) rs71781329 konnte in der Detektionsstichprobe eine vollständige Übertragung des Minorallels auf das betroffene Kind nachgewiesen werden, welche eine signifikante Assoziation mit ASS zeigte (p-Wert = 0,041). Diese Beobachtung konnte jedoch weder in der Replikationsstichprobe, noch im Gesamtset repliziert werden. Vermutlich reichte die Stichprobengröße aufgrund der geringen Allelfrequenz bei weitem nicht aus, um das Ergebnis replizieren zu können. Deskriptiv zeigte sich aber in der Gesamtstichprobe, dass das Minorallel GCG[7] untertransmittiert vorliegt während für das Minorallel GCG[8] hingegen eine Übertransmission beobachtet werden konnte. Diesem Ergebnis könnte eine allelspezifische Vererbung der Repeats in ASS zugrunde liegen.

Bei der Regressionsanalyse zeigte sich, dass zwischen heterozygoten Trägern des Allels GCG[7] und homozygoten Majorallelträgern signifikante Unterschiede in der Sprachentwicklung vorliegen. So sprechen heterozygote Minorallelträger ihre ersten Worte durchschnittlich 53 Monate später (SD = 42,43; p-Wert < 0,0001) und ihre ersten Sätze durchschnittlich 62 Monate (SD = 38,18; p-Wert = 0,0001) später als die homozygoten Majorallelträger (GCG[6]). Interessanterweise wurde für heterozygote Träger des Allels GCG[8] der gleichen Variante kein Unterschied zu den homozygoten GCG[6] Trägern in der Sprachentwicklung nachgewiesen. Da jedoch für das Allel GCG[7] des STR nur zwei Patienten und für Allel GCG[8] nur drei Patienten in unserer Gesamtstichprobe nachgewiesen werden konnten, müssen diese Ergebnisse mit Vorsicht interpretiert werden.

Sollten diese Ergebnisse jedoch in weiteren Studien bestätigt werden, so kann angenommen werden, dass die vorliegenden Mikrosatelliten die *CNTNAP2*-Expression in Abhängigkeit der Längenunterschiede beeinflussen und so auf unterschiedliche, spezifische Weise auf den Sprachphänotypen einwirken.

7.4 Transfektionseffizienztest

Es ist allgemein bekannt, dass die Transfektion der Neuroblastomzelllinie SH-SY5Y mit erheblichen Limitierungen verbunden ist. So wird auch in der Literatur die Transfektionseffizienz für einige an

dieser Zelllinie getesteten Transfektionsreagenzien mit gerade Mal 5% angegeben (Martín-Montañez et al. 2010).

Bei der hier durchgeführten Optimierung der Transfektionseffizienz zeigte sich, dass das Transfektionsreagenz Metafectene® Pro (Biontex Laboratories GmbH) bereits nach 24 Stunden mit fast 30% und nach 48 Stunden mit fast 50% die höchste Transfektionseffizienz in den Zellen der SH-SY5Y Zelllinie aufweist. Bei anderen Zelllinien, wie beispielsweise der ebenfalls im Rahmen dieser Studie eingesetzten HEK293T-Zelllinie, können zum Teil Effizienzen von 90-100% erzielt werden. Aufgrund der Schwierigkeiten bei der Transfektion der SH-SY5Y Zelllinie gelten 30-50% bereits als ein sehr gutes Ergebnis. Durch weitere Optimierungsversuche auf Ebene der Zelldichte und Mediumkomposition kann unter Umständen eine noch etwas höhere Transfektionseffizienz erzielt werden. Der hier durchgeführte Transfektionseffizienztest zielte jedoch einzig und allein auf die Bestimmung des Reagenzes mit der höchsten Transfektionseffizienz ab und da die Ergebnisse des Tests sehr eindeutig waren wurde aus Zeit- und Kostengründen von der exakten Bestimmung der Transfektionseffizienz und weiteren Transfektionsoptimierungen abgesehen.

Größere Schwierigkeiten ergaben sich bei dem Versuch, die bereits differenzierten Neuroblastomzellen zu transfizieren. Selbst mit dem Transfektionsreagenz Metafectene® Pro (Biontex Laboratories GmbH) lag die Transfektionseffizienz der differenzierten Zellen bei < 1%. Dieses Problem konnte umgangen werden, indem mit der Differenzierung der Zellen erst nach der Transfektion der Zellen begonnen wurde. Dadurch lag die Transfektionseffizienz deutlich höher (~40%). Normalerweise ist bei einer transienten Transfektion, wie sie hier durchgeführt wurde, die Lebensdauer der aufgenommenen DNA auf nur wenige Tage beschränkt (Mülhardt 2006), da fremde DNA in der Zelle entweder abgebaut, oder bei der mitotischen Teilung nicht mitgeschrieben und weitergegeben wird. Die Kultivierung der Zellen über einen so langen Zeitraum und der Erhalt der DNA war im Rahmen dieser Studie möglich, da sich differenzierende Zellen nicht teilen und das aufgenommene Plasmid somit über den gesamten Zeitraum in den Zellen verblieben ist. Für zukünftige Versuche sollte jedoch darüber nachgedacht werden, eine stabile Transfektion der Zellen durchzuführen, bei der die Zellen die Plasmid-DNA genomisch integrieren, da so gewährleistet werden kann, dass zu jedem Messzeitpunkt die Zielsequenz in den Zellen vorliegt. Hierbei muss jedoch bedacht werden, dass eine stabile Transfektion nicht nur Vorteile mit sich bringt, da durch die genomische Integration der Plasmide für die geplante Untersuchung relevante Gensequenzen unterbrochen bzw. gestört werden können. Zudem sind die Zellen bei einer stabilen

Transfektion durch den erhöhten Arbeitsaufwand nicht wie transient transfizierte Zellen bereits nach 24-48 Stunden für Experimente einsetzbar, sondern erst nach etwa 2 Monaten. Ebenso unterliegen stabil transfizierte Zelllinien einem erhöhten Selektionsdruck und bringen aufgrund der langen Zeit in Kultur eine erhöhte Mutationsgefahr mit sich (Mülhardt 2006).

Bezüglich der Luciferaseaktivitätsmessung stellt die Kultivierung der Zellen über einen längeren Zeitraum prinzipiell kein Problem dar, da die Halbwertszeit der verwendeten Luciferase bei ~3 Stunden liegt (Leclerc et al. 2000; Thompson et al. 1991) und man somit auch nach mehrtägiger Differenzierung der Zellen einen Schnappschuss der jeweiligen Promotoraktivität erhält. Dennoch sollte man nicht außer Acht lassen, dass die Halbwertszeit der gebildeten mRNA bei ~6 Stunden liegt (Thompson et al. 1991) und sie somit doppelt so lang stabil bleibt als das gebildete Protein. Dadurch könnte es zu einer Verfälschung der Werte der eigentlichen Promotoraktivität kommen, insbesondere wenn zu einem bestimmten Zeitpunkt eine Runterregulierung des Promotors zu erwarten ist. Da in dieser Arbeit das Zeitfenster zwischen den einzelnen Messungen jedoch größer als 6 Stunden war, kann davon ausgegangen werden, dass die beobachteten Effekte nicht durch die Stabilität der Luciferase-Zwischenprodukte (mRNA bzw. Protein) beeinflusst wurden.

7.5 Funktionelle Analyse der Varianten

Ein weiteres zentrales Augenmerk dieser Studie lag auf der Untersuchung der Funktion der *CNTNAP2*-Promotorvarianten und deren Einfluss auf die Genexpression während der neuronalen Differenzierung. Die in-silico Analyse hat gezeigt, dass durch die verschiedenen Allele der Varianten die Bindestellen von Transkriptionsfaktoren verändert werden. In der in-vitro Analyse konnte letztendlich bestätigt werden, dass die Varianten die Promotoraktivität, abhängig von dem zellulären Hintergrund und dem Differenzierungsstadium der Zellen, unterschiedlich beeinflussen.

Bei der in-silico Analyse zeigte sich, dass durch das Minor-G-Allel des nicht mit ASS assoziierten SNP rs150447075 eine zusätzliche Bindestelle für den hirnspezifischen, onkogenen Krüppel like Transkriptionsfaktor GLI3 (Glioma-associated oncogene family zinc finger 3) entsteht, welcher an der Extremitätenausbildung beteiligt ist (Büscher et al. 1997). Dieser Transkriptionsfaktor ist vor allem in mitotischen Zellen aktiv und reguliert die Zellteilung und Apoptose während der Embryogenese. Eine Promotorvariante kann jedoch vermutlich nur einen Effekt auf die Genexpression ausüben, wenn der interagierende Transkriptionsfaktor auch verfügbar ist. Da für GLI3 kein gehirnspezifischer Effekt

bekannt ist, könnte dies erklären, warum diese Variante in der vorliegenden Assoziationsstudie keinen Einfluss auf das ASS Risiko bzw. auf die Sprachentwicklung hat. Bei der in-vitro Analyse zeigte sich zudem eine durch diese Variante hervorgerufene signifikant gesteigerte Transkriptionseffizienz in den mitotischen Zellen der Zelllinien HEK293T (p-Wert 0,050) und SH-SY5Y (p-Wert 0,008), die jedoch weder 72 Stunden (p-Wert 0,234), noch 216 Stunden (p-Wert 0,455) nach Induktion der Differenzierung in den SH-SY5Y Zellen nachgewiesen werden konnte. Dies deckt sich mit der Erwartung, dass GLI3 vor allem in mitotischen Zellen eine gesteigerte Aktivität zeigt, was wiederum einen Hinweis darauf gibt, dass in den kultivierten Zellen die Bindestelle für diesen TF wirklich durch das G-Allel der Variante rs150447075 verändert wird. Die hier erhaltenen Ergebnisse zeigen aber, dass diese Variante trotz des funktionellen Effekts nicht das Risiko für ASS beeinflusst, was sich vermutlich durch erklären lässt, dass es sich bei ASS um eine neurologische Erkrankung handelt und für GLI3 keine gehirnspezifischen Effekte bekannt sind.

Durch das Minor-G-Allel der Variante rs34712024 fallen bei der in-silico Analyse die Bindestellen für die Transkriptionsfaktoren ZF300, TFAP4 und Myogenin (MYOG) weg. ZF300 ist ein Protein, das in Herz, Skelettmuskulatur und Gehirn exprimiert wird und dort überwiegend im Nukleus vorzufinden ist. Auch wenn seine Funktion bisher nicht vollends geklärt werden konnte, wird aufgrund seiner terminalen KRAB-Domäne vermutet, dass er als Repressor der Transkription fungiert (Gou et al. 2004). TFAP4 ist ein bifunktionaler Transkriptionsfaktor, der auch im Gehirn exprimiert wird und sowohl als Aktivator, als auch als Repressor auf die jeweiligen Zielgene einwirken kann. Des Weiteren konnte gezeigt werden, dass TFAP4 bei der Transkription nach DNA-Schädigungen eine große Rolle spielt (Ku et al. 2009). Myogenin (MYOG) hingegen ist ein muskelspezifischer Transkriptionsfaktor, der für die Einleitung der Myogenese verantwortlich ist (Chen et al. 2011). Durch das Allel rs34712024G entsteht noch eine zusätzliche Bindestelle für den Krüppel like Transkriptionsfaktor (TF) ZF9, der während der Zelldifferenzierung und dem Zellwachstum als starker Transkriptionenhancer agiert (DiFeo et al. 2009). Dass durch den TF MYOG die von uns in der Assoziationsstudie beobachteten Effekte hervorgerufen wurden ist aufgrund seiner beschriebenen Funktion eher auszuschließen. Wesentlich wahrscheinlicher ist es, dass durch die veränderten Bindestellen der hirnspezifischen Transkriptionsfaktoren eine Verschiebung in der Regulierung der Genexpression entsteht und somit das ASS Risiko bzw. der Sprachphänotyp in ASS beeinflusst werden. Bei der in-vitro Analyse zeigte sich in den mitotischen Zellen der Zelllinie HEK293T eine durch diese Variante hervorgerufene signifikant gesteigerte Transkriptionseffizienz

(p-Wert = 0,038). In den mitotischen Zellen der Neuroblastomzelllinie konnte hingegen eine leichte Runterregulierung der Transkription beobachtet werden, die jedoch nicht signifikant (p-Wert = 0,105) wurde. Auch zu den übrigen Zeitpunkten während der Differenzierung konnten keine signifikanten Änderungen der Transkriptionseffizienz beobachtet werden. Man kann aufgrund der vorliegenden Beobachtungen nun vermuten, dass sowohl der protektive Effekt auf das ASS-Risiko, als auch der negative Effekt auf den Sprachphänotypen auf die veränderten Bindestellen der Transkriptionsfaktoren TFAP4 und ZF9 zurückzuführen sind. So könnten diese Transkriptionsfaktoren aufgrund ihrer unterschiedlichen aktivierenden und hemmenden Rolle während der neuronalen Entwicklung die Expression von *CNTNAP2* zu unterschiedlichen Zeitpunkten in entgegengesetzten Richtungen beeinflussen, was die an sich widersprüchlichen Effekten erklären könnte. Dass in unserem in-vitro Versuch während der Differenzierung der SH-SY5Y Zellen keine regulatorischen Vorgänge zu beobachten sind liegt vermutlich daran, dass die in-vivo vorliegenden genetischen oder umweltbedingten Risikofaktoren, sowie eventuelle Interaktionen mit Transkriptionsfaktoren auf 3' liegenden Sequenzen des Gens, in dem Versuchsaufbau nicht berücksichtigt werden konnten.

Bei der in-silico Analyse des STR rs71781329 zeigte sich, dass auch die Minorallele GCG[7] und GCG[8] als Modulatoren der Transkription agieren können, da pro eingefügtem Repeat eine zusätzliche Bindestelle für das „Early growth response protein 1“ (EGR1) entsteht, welches eine zentrale Rolle in der Regulierung der neuronalen Plastizität und Differenzierung spielt (Jones et al. 2001; Lee et al. 2004). Zudem wird die Expression von EGR1 durch synaptische Aktivität, einem Mechanismus, der bei ASS vermutlich beeinträchtigt ist (Ebert und Greenberg 2013), reguliert. Abweichungen in der EGR1 Expression beeinflussen wiederum die Transkription der Zielgene, inklusive *CNTNAP2* (Knapska und Kacmarek 2004). Des Weiteren hat sich gezeigt, dass bei EGR1-defizienten Mäusen die Überführung von Lernprozessen ins Langzeitgedächtnis gehemmt ist (Jones et al. 2001). Im Zelldifferenzierungsmodell der SH-SY5Y Zellen zeigt interessanterweise GCG[7] und nicht GCG[8] einen Trend zur differentiellen Regulierung in den undifferenzierten Zellen (p-Wert GCG[7] = 0,058) mit einer normalen Expression in den späteren Differenzierungszeitpunkten (nach 72 und 216 Stunden). Für GCG[8] ist hingegen erst während der Differenzierungsphase (nach 72 Stunden) ein signifikanter (p-Wert = 0,022), regulierender Effekt aufgetreten. Dass Längenunterschiede von Mikrosatelliten im Promotor die Genexpressionsrate signifikant ändern und so die Hirnfunktion und -plastizität interindividuell beeinflussen können, wurde bereits für das PAX-6 Gen

gezeigt (Okladnova et al. 1998). Zudem ist bekannt, dass Mikrosatelliten in der Promotorregion Sekundärstrukturen in der DNA ausbilden können, durch die wiederum eine Regulation der Genexpression erfolgt (Kouzine 2007; Sawaya et al. 2012). Es ist daher zusätzlich anzunehmen, dass die Genexpressionsrate durch die Ausbildung von Sekundärstrukturen der verschiedenen Allele des STR rs71781329 beeinflusst wird.

Die Varianten, für die im Rahmen dieser Studie eine Assoziation mit ASS und dem Sprachphänotypen nachgewiesen werden konnten ändern interessanterweise Bindesequenzen von Transkriptionsfaktoren, die auch im Gehirn eine zentrale Rolle spielen. Die Effekte könnten somit auf die modifizierten Transkriptionsfaktorbindestellen zurückgeführt werden. Es wird in der Literatur beschrieben, dass der Transkriptionsfaktor FOXP2 die *CNTNAP2*-Expression in Föten der 18-22 Gestationswoche deutlich herunterreguliert (Vernes et al. 2008) und als Hauptmodulator bei der *CNTNAP2*-Expression agiert. Die in den Föten nachgewiesene Regulation der *CNTNAP2*-Expression deckt sich mit den Ergebnissen der mRNA Expressionsstudie, die Frau Denise Haslinger (M.Sc.) in unserem Labor durchgeführt hat (Daten sind bisher nicht publiziert). Hierbei konnte gezeigt werden, dass *CNTNAP2* 72 Stunden nach Induktion der Zelldifferenzierung (pIZ) in der Neuroblastomzelllinie SH-SY5Y stark herunterreguliert wird. Dies entspricht auch dem Ergebnis einer anderen Studie mit humanen neuronalen Vorläuferzellen, in der eine Runterregulierung von *CNTNAP2* während der neuronalen Differenzierung nachgewiesen werden konnte (Konopka et al. 2012). In einer Studie von Kang und Kollegen (2011) zeigte sich, dass in gesunden Individuen im Gehirn die *CNTNAP2*-Expression prä- und perinatal in allen Regionen erfolgt, sie jedoch im frühen Kindesalter (1-6 Jahren) stark runterreguliert wird. Daran anschließend erfolgt eine lebenslange Re-expression von *CNTNAP2* in allen Hirnregionen, mit Ausnahme des Striatum und des Cerebellar Cortex, zwei Bereiche, deren Beteiligung an ASS (Haas et al. 1996; Langen et al. 2009) und am hierarchischen Lernen während des Spracherwerbs bereits diskutiert wurde (Strick et al. 2009; Desrochers und Badre 2012). Auch im Rahmen dieser Studie zeigte sich eine deutliche Erhöhung der Expressionsaktivität des *CNTNAP2*-Promotors 216 Stunden pIZ, wodurch die Aktivierung des Promotors definitiv belegt werden konnte. Es lässt sich daher vermuten, dass die Runterregulierung der *CNTNAP2*-Expression, die in der hier durchgeführten Promotoraktivitätsstudie nicht nachweisbar war, hauptsächlich über FOXP2 stattfindet. Die Bindestelle für FOXP2 liegt jedoch in der intronischen Region zwischen Exon 1 und 2 (Vernes et al. 2008), sodass es nahe liegt zu vermuten, dass die Inhibition während der neuronalen Differenzierung über FOXP2 gesteuerte Mechanismen,

die Re-Aktivierung des *CNTNAP2*-Promotors aber *FOXP2* unabhängig über regulatorische Elemente in der hier untersuchten 5' gelegenen Promotorregion erfolgt. Die hier analysierten Varianten zeigten einen Effekt vor allem zu Zeitpunkten an denen keine erhöhte Aktivierung des Promotors beobachtet wurde, sodass angenommen werden kann, dass bereits in den inhibitorischen Phasen der pathologische Effekt der Varianten zum Tragen kommt. Aufgrund der hohen Expressionsrate zum Zeitpunkt 216 Stunden pIZ und den damit verbundenen erhöhten Schwankungen in den Messwerten sind allerdings geringfügige Änderungen in der Promotoraktivität, wie sie für diese Varianten ebenso zu erwarten gewesen wären, zu diesem Zeitpunkt nur schwer nachweisbar. Zudem kann angenommen werden, dass die Regulationen durch *FOXP2* geringfügige Modulationen anderer Transkriptionsfaktoren ausgleichen und somit überdecken. Das könnte auch der Grund sein, warum die hier beschriebenen Promotorvarianten bisher nicht mit der Regulierung von *CNTNAP2* in Verbindung gebracht werden konnten.

Es wäre zudem denkbar, dass es durch die veränderten Transkriptionsfaktorbindestellen zu der hier beobachteten, verschobenen Feinabstimmung in der Transkription von *CNTNAP2* an bestimmten Zeitpunkten der neuronalen Entwicklung kommt. Dies könnte wiederum einen spezifischen Effekt auf den Sprachphänotypen in ASS haben, wobei insbesondere die reduzierte *CNTNAP2*-Expression in einem sehr frühen, nicht aber in einem späteren Stadium der Differenzierung einen Einfluss auf den ASS Phänotypen zu haben scheint.

7.6 Limitierung der Studie

Die Limitierung der Studie erfolgte durch mehrere Faktoren. Einer der wichtigsten Punkte ist hierbei die Power der Stichproben. Es ist erstaunlich, dass trotz der seltenen Allelfrequenz und der Verhältnismäßig kleinen Stichprobe bereits in der Detektionsstichprobe signifikante Ergebnisse erzielt wurden. Ebenso muss erwähnt werden, dass die p-values der ORs das Signifikanzniveau von $\alpha = 0,05$ nach Korrektur für multiples Testen (p-Wert (Bonferroni) = 0,0012) nicht unterschritten haben.

Ein weiterer wichtiger Punkt ist, dass bei einigen Patienten keine Daten des Diagnostischen Interviews für Autismus - Revidiert (ADI-R; Lord et al. 1994) vorliegen, da die Diagnosestellung ausschließlich aufgrund der Beobachtungsskala für Autistische Störungen (ADOS-G; Lord et al. 2000) erfolgte. So sind durch die fehlenden Werte für das Alter bei dem die ersten Worte gesprochen wurden (ADI_A9) nur 65%, für das Alter in dem die ersten Sätze gesprochen wurden

(ADI_A10) sogar nur 57% der Patienten in die Regressionsanalyse mit eingeflossen. Auch wurde nicht bei allen Patienten der IQ bestimmt und zum Teil fehlen die Angaben zum Alter bei Diagnose, was zu einer weiteren Reduktion der Stichprobe und der damit verbundenen statistischen Power in der Regressionsanalyse führte.

Ebenso sollte die retrospektive Durchführung der Elterninterviews kritisch betrachtet werden. Es stellt sich nämlich die Frage, inwieweit die Angaben der Eltern, deren Kinder bei der Diagnosestellung bereits älter waren, durch Lücken und Fehler in der Erinnerung beeinträchtigt werden können. Es hat sich gezeigt, dass es von der Art der retrospektiven Frage abhängt, ob ein Ereignis richtig datiert und erinnert wird, unabhängig davon wie weit es zurückliegt. So werden Ereignisse, die einen zentralen lebenszyklischen Wendepunkt darstellen, im Allgemeinen sehr zuverlässig erinnert und werden hierbei auch kaum durch fehlende Reliabilität beeinträchtigt (Klein und Fischer-Kerli 2000). Schwieriger zu erfassen sind hingegen Ereignisse, die sich im Verlaufe eines Lebens mehrfach wiederholen, da die Gefahr groß ist, dass vor allem Ereignisse die weit zurückliegen miteinander verwechselt werden bzw. die Erinnerungen daran verblassen (Lewandowsky und Murdock Jr. 1989). Da es sich in den Elterninterviews jedoch meist um einmalige und wichtige Ereignisse im Leben des Kindes handelt, wie z.B. das Sprechen erster Worte, kann man davon ausgehen, dass die Daten richtig erinnert werden. Allerdings ist eine falsche Zuordnung der Daten nicht vollständig auszuschließen, weswegen das Alter des Kindes zum Zeitpunkt der Diagnosestellung bei der Regressionsanalyse auch als Kovariante in das Modell mit eingeflossen ist.

Der letzte Punkt, der zur Limitierung der Studie beiträgt ist die Transfektionseffizienz bei den Zellen der SH-SY5Y Zelllinie. Auch nach der durchgeführten Optimierung ist die Transfektionseffizienz für diese Zelllinie deutlich geringer als für die Zellen der HEK293T Zelllinie, die bei fast allen Transfektionsreagenzien und -bedingungen eine Transfektionseffizienz von > 90% aufweisen. Es ist daher möglich, dass sowohl die geringe Signalausbeute, als auch die Streuung der einzelnen Messungen durch die geringe Transfektionseffizienz beeinflusst werden und somit auch den Effekt der Allele verfälscht bzw. maskiert.

7.7 Ausblick

Für den SNP rs34712024 wäre es wichtig eine weitere Genotypisierung in einer unabhängigen und phänotypisch gut charakterisierten Stichprobe durchzuführen, um die Ergebnisse dieser Studie zu

validieren. Des Weiteren sollte der modulierende Effekt, den das Minorallel rs71781329GCG[7] auf den Sprachphänotypen ausübt genauer untersucht werden. Hierbei sollte die Replikation der Ergebnisse in einer unabhängigen Stichprobe angestrebt werden. Diese muss jedoch ausreichend groß sein, um die beiden Allele der Variante getrennt von einander untersuchen zu können.

Um die pleiotropen und multidirektionalen Effekte der Varianten auf das ASS Risiko und den Sprachphänotypen aufklären zu können sollten bei einer Replikation der vorliegenden Ergebnisse die Analysen sowohl in einer unabhängigen Stichprobe mit autistischen Kindern, als auch in einer Stichprobe mit Kindern, die eine spezifische Sprachstörung (SSES) aufweisen, sowie in einer gesunden Kontrollgruppe durchgeführt werden. Nur so ist es möglich die Krankheitsbilder direkt gegeneinander zu vergleichen und die Effekte spezifischen Phänotypen zuzuordnen. Letztendlich ist es wichtig den Effekt der Varianten in Bezug auf Krankheitsbild und Phänotyp abzugrenzen.

Um die genetische Befunde zellbiologischen Vorgängen zuordnen zu können, bedarf es zudem einer funktionellen Analyse. Die hier berichteten zelltyp- und differenzierungsstatusspezifischen Effekte von Varianten zeigen, dass die pathologischen Mechanismen abhängig vom untersuchten System sind, was in zukünftigen funktionellen Studien auch berücksichtigt werden müsste.

Ein Hauptziel zukünftiger Studien sollte daher sein, die Grundlagen der Effekte auch in Krankheitsbildern ohne ASS Symptomatik, aber mit sprachbezogenen Phänotypen wie zum Beispiel SSES, Dyslexie oder dem Gilles-de-la-Tourette-Syndrom genetisch und funktionell zu untersuchen, um letztendlich die Unterschiede in der Ätiologie von Sprachphänotypen und ASS aufzuklären und besser verstehen zu können.

8 Zusammenfassung und Schlussfolgerung

Im Rahmen dieser Studie wurde eine genetische Analyse der drei häufigsten Varianten der *CNTNAP2*-Promotorregion mit einer anschließenden funktionellen Untersuchung durchgeführt. Als Studienmaterial diente die DNA einer phänotypisch gut charakterisierten Autismusstichprobe. Für das Minor-G-Allel der Variante rs34712024 konnte hierbei einerseits ein protektiver Effekt auf das ASS-Risiko, andererseits ein negativer Einfluss auf den Sprachphänotypen festgestellt werden. Dieser bidirektionale Effekt ist möglicherweise auf die durch die Minorallele veränderte Transkriptionseffizienz des Promotors zurückzuführen, welche durch die veränderten Transkriptionsfaktorbindestellen vermittelt werden könnte.

Im Rahmen dieser Studie konnte die Frage nach der Ursache für ASS sowohl auf genetischer als auch auf funktioneller Ebene weiter geklärt werden. Es hat sich deutlich gezeigt, dass nicht nur das Screening nach krankheitsverursachenden genetischen Varianten von Belang ist, sondern es auch von enormer Bedeutung ist, diese in einer anschließenden funktionellen Studie weiter zu untersuchen, um diese mit zellulären Mechanismen in Bezug zu bringen. Denn die Betrachtung komplexer pathomechanischer Zusammenhänge kann letztendlich zur Aufklärung der Ätiologie von ASS einen enormen Beitrag leisten

Diese Arbeit wurde am 18. Juni 2013 bei „Molecular Psychiatry“ eingereicht und befindet sich zum Zeitpunkt der Abgabe in Revision. Eine Kopie des eingereichten Manuskripts ist der Arbeit hinten beigelegt.

9 Literatur

- Abrahams, B. S.; Tentler, D.; Perederiy, J. V.; Oldham, M. C.; Coppola, G.; Geschwind, D. H. (2007): Genome-wide analyses of human perisylvian cerebral cortical patterning. In: *Proceedings of the National Academy of Sciences* 104 (45), S. 17849–17854.
- Abrahams, Brett S.; Geschwind, Daniel H. (2008): Advances in autism genetics: on the threshold of a new neurobiology. In: *Nat Rev Genet* 9 (5), S. 341–355.
- Alarcón, Maricela; Abrahams, Brett S.; Stone, Jennifer L.; Duvall, Jacqueline A.; Perederiy, Julia V.; Bomar, Jamee M. et al. (2008): Linkage, Association, and Gene-Expression Analyses Identify CNTNAP2 as an Autism-Susceptibility Gene. In: *The American Journal of Human Genetics* 82 (1), S. 150–159.
- American Psychiatric Association (Hg.) (2000): Diagnostic and statistical manual of mental disorders. DSM-IV-TR. 4. Aufl. Washington, DC: American Psychiatric Association.
- Arking, Dan E.; Cutler, David J.; Brune, Camille W.; Teslovich, Tanya M.; West, Kristen; Ikeda, Morna et al. (2008): A Common Genetic Variant in the Neurexin Superfamily Member CNTNAP2 Increases Familial Risk of Autism. In: *The American Journal of Human Genetics* 82 (1), S. 160–164.
- Asperger, H. (1944): Die autistischen Psychopathen im Kindesalter. In: *Archiv für Psychiatrie und Nervenkrankheiten* (177), S. 76–166.

- Bailey, A.; Palferman, S.; Heavey, L.; Le Couteur, A. (1998): Autism: the phenotype in relatives. In: *J Autism Dev Disord* 28 (5), S. 369–392.
- Bakkaloglu, Betul; O'Roak, Brian J.; Louvi, Angeliki; Gupta, Abha R.; Abelson, Jesse F.; Morgan, Thomas M. et al. (2008): Molecular Cytogenetic Analysis and Resequencing of Contactin Associated Protein-Like 2 in Autism Spectrum Disorders. In: *The American Journal of Human Genetics* 82 (1), S. 165–173.
- Barrett, J. C.; Fry, B.; Maller, J.; Daly, M. J. (2005): Haploview: analysis and visualization of LD and haplotype maps. In: *Bioinformatics* 21 (2), S. 263–265.
- Bisgaard, Anne-Marie; Rackauskaite, Gija; Thelle, Thomas; Kirchhoff, Maria; Bryndorf, Thue (2006): Twins with mental retardation and an interstitial deletion 7q34q36.2 leading to the diagnosis of long QT syndrome. In: *Am. J. Med. Genet. A* 140 (6), S. 644–648.
- Bishop, Dorothy V. M.; Maybery, Murray; Maley, Alana; Wong, Dana; Hill, Wayne; Hallmayer, Joachim (2004): Using self-report to identify the broad phenotype in parents of children with autistic spectrum disorders: a study using the Autism-Spectrum Quotient. In: *J Child Psychol Psychiatry* 45 (8), S. 1431–1436.
- Bleuler, Eugen (Hg.) (1911): *Dementia praecox oder Gruppe der Schizophrenien*. Tübingen: Edition Diskord.
- Bölte, Sven (Hg.) (2009): *Autismus. Spektrum, Ursachen, Diagnostik, Intervention, Perspektiven*. 1. Aufl. Bern: Huber (Psychologie Lehrbuch).
- Bölte, Sven; Knecht, Susan; Poustka, Fritz (2007): A case-control study of personality style and psychopathology in parents of subjects with autism. In: *J Autism Dev Disord* 37 (2), S. 243–250.
- Bonifazi, P.; Goldin, M.; Picardo, M. A.; Jorquera, I.; Cattani, A.; Bianconi, G. et al. (2009): GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. In: *Science* 326 (5958), S. 1419–1424.
- Burgemeister, Bessie B.; Hollander Blum, Lucille; Lorge, Irving (1954): *The Columbia Mental Maturity Scale*. Yonkers-Hudson.
- Büscher, Dirk; Bosse, Birgit; Heymer, Joachim; Rüther, Ulrich (1997): Evidence for genetic control of Sonic hedgehog by Gli3 in mouse limb development. In: *Mechanisms of Development* 62 (2), S. 175–182.
- Campbell, Neil A. (1997): *Biologie*. Heidelberg, Berlin: Spektrum Akademie. (Spektrum-Lehrbuch).
- Caselli, Rossella; Mencarelli, Maria Antonietta; Papa, Filomena Tiziana; Ariani, Francesca; Longo, Ilaria; Meloni, Ilaria et al. (2008): Delineation of the phenotype associated with 7q36.1q36.2 deletion: long QT syndrome, renal hypoplasia and mental retardation. In: *Am. J. Med. Genet. A* 146 (9), S. 1195–1199.

- Chen, L.; Huang, H. W.; Gu, S. H.; Xu, L.; Gu, Y. D.; Xu, J. G. (2011): The study of myogenin expression in denervated human skeletal muscles. In: *J. Int. Med. Res.* 39 (2), S. 378–387.
- Chiocchetti, A. (2012): Behind the autism spectrum. In: *Science in school* 24 (Autumn), S. 17-22.
- Courchesne, Eric; Pierce, Karen; Schumann, Cynthia M.; Redcay, Elizabeth; Buckwalter, Joseph A.; Kennedy, Daniel P.; Morgan, John (2007): Mapping early brain development in autism. In: *Neuron* 56 (2), S. 399–413.
- Curran, Sarah; Bolton, Patrick; Rozsnyai, Kinga; Chiocchetti, Andreas; Klauck, Sabine M.; Duketis, Eftichia et al. (2011): No association between a common single nucleotide polymorphism, rs4141463, in the MACROD2 gene and autism spectrum disorder. In: *Am. J. Med. Genet.* 156 (6), S. 633–639.
- Desrochers, Theresa M.; Badre, David (2012): Finding parallels in fronto-striatal organization. In: *Trends Cogn. Sci. (Regul. Ed.)* 16 (8), S. 407–408.
- DiFeo, Analisa; Narla, Goutham; Martignetti, John A. (2009): Emerging roles of Kruppel-like factor 6 and Kruppel-like factor 6 splice variant 1 in ovarian cancer progression and treatment. In: *Mt. Sinai J. Med.* 76 (6), S. 557–566.
- Doja, Asif; Roberts, Wendy (2006): Immunizations and autism: a review of the literature. In: *Can J Neurol Sci* 33 (4), S. 341–346.
- Dosenbach, N. U. F.; Nardos, B.; Cohen, A. L.; Fair, D. A.; Power, J. D.; Church, J. A. et al. (2010): Prediction of Individual Brain Maturity Using fMRI. In: *Science* 329 (5997), S. 1358–1361.
- Dudbridge, Frank (2008): Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. In: *Hum. Hered.* 66 (2), S. 87–98.
- Ebert, Daniel H.; Greenberg, Michael E. (2013): Activity-dependent neuronal signalling and autism spectrum disorder. In: *Nature* 493 (7432), S. 327–337.
- Elsabbagh, Mayada; Divan, Gauri; Koh, Yun-Joo; Kim, Young Shin; Kauchali, Shuaib; Marcín, Carlos et al. (2012): Global prevalence of autism and other pervasive developmental disorders. In: *Autism Res* 5 (3), S. 160–179.
- Encinas, M.; Iglesias, M.; Liu, Y.; Wang, H.; Muhaisen, A.; Ceña, V. et al. (2000): Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. In: *J. Neurochem.* 75 (3), S. 991–1003.
- Folstein, S.; Rutter, M. (1977): Infantile autism: a genetic study of 21 twin pairs. In: *J Child Psychol Psychiatry* 18 (4), S. 297–321.
- Fombonne, E. (2002): Epidemiological trends in rates of autism. In: *Mol Psychiatry* 7 (s2), S. S4.
- Fombonne, Eric (2003): Epidemiological surveys of autism and other pervasive developmental disorders: an update. In: *J Autism Dev Disord* 33 (4), S. 365–382.

- Fombonne, Eric (2005): Epidemiology of autistic disorder and other pervasive developmental disorders. In: *J Clin Psychiatry* 66 Suppl 10, S. 3–8.
- Freitag, Christine M.; Staal, Wouter; Klauck, Sabine M.; Duketis, Eftichia; Waltes, Regina (2010): Genetics of autistic disorders: review and clinical implications. In: *Eur Child Adolesc Psychiatry* 19 (3), S. 169–178.
- Freitag, Christine M. (2012): Autistische Störungen – State-of-the-Art und neuere Entwicklungen. In: *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 40 (3), S. 139–149.
- Gauderman W, Morrison J.: QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. Online verfügbar unter <http://hydra.usc.edu/gxe>.
- Geschwind, Daniel H. (2008): Autism: Many Genes, Common Pathways? In: *Cell* 135 (3), S. 391–395.
- Geschwind, Daniel H.; Levitt, Pat (2007): Autism spectrum disorders: developmental disconnection syndromes. In: *Curr. Opin. Neurobiol.* 17 (1), S. 103–111.
- Gillberg, Christopher; Cederlund, Mats; Lamberg, Kerstin; Zeijlson, Lena (2006): Brief Report: “The Autism Epidemic”. The Registered Prevalence of Autism in a Swedish Urban Area. In: *J Autism Dev Disord* 36 (3), S. 429–435.
- Gou, Deming; Wang, Jun; Gao, Li; Sun, Yan; Peng, Xiao; Huang, Jian; Li, Wenxin (2004): Identification and functional analysis of a novel human KRAB/C2H2 zinc finger gene ZNF300. In: *Biochim. Biophys. Acta* 1676 (2), S. 203–209.
- Guerrini, Renzo; Barba, Carmen (2010): Malformations of cortical development and aberrant cortical networks: epileptogenesis and functional organization. In: *J Clin Neurophysiol* 27 (6), S. 372–379.
- Haas, R. H.; Townsend, J.; Courchesne, E.; Lincoln, A. J.; Schreibman, L.; Yeung-Courchesne, R. (1996): Neurologic abnormalities in infantile autism. In: *J. Child Neurol.* 11 (2), S. 84–92.
- Haesler, Sebastian; Rochefort, Christelle; Georgi, Benjamin; Licznarski, Pawel; Osten, Pavel; Scharff, Constance (2007): Incomplete and Inaccurate Vocal Imitation after Knockdown of FoxP2 in Songbird Basal Ganglia Nucleus Area X. In: *Plos Biol* 5 (12), S. e321.
- Hallmayer, Joachim; Cleveland, Sue; Torres, Andrea; Phillips, Jennifer; Cohen, Brianne; Torigoe, Tiffany et al. (2011): Genetic heritability and shared environmental factors among twin pairs with autism. In: *Arch. Gen. Psychiatry* 68 (11), S. 1095–1102.
- Holtermann, M.; Bölte, S.; Poustka, F. (2006): Genetik des Autismus. In: *Medizinische Genetik* (2, 18, 2), S. 170–174.
- International HapMap Consortium (2005): A haplotype map of the human genome. In: *Nature* 437 (7063), S. 1299–1320.

- International Molecular Genetic Study of Autism Consortium (1998): A full genome screen for autism with evidence for linkage to a region on chromosome 7q. International Molecular Genetic Study of Autism Consortium. In: *Hum. Mol. Genet.* 7 (3), S. 571–578.
- International Molecular Genetic Study of Autism Consortium (2001): A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. In: *Am. J. Hum. Genet.* 69 (3), S. 570–581.
- Jones, M. W.; Errington, M. L.; French, P. J.; Fine, A.; Bliss, T. V. P.; Garel, S. et al. (2001): A requirement for the immediate early gene *Zif268* in the expression of the late LTP and long-term memories. In: *Nat. Neurosci.* 4 (3), S. 289–296.
- Kang, Hyo Jung; Kawasaki, Yuka; Imamura, Cheng, Feng; Zhu, Ying; Xu, Xuming; Li, Mingfeng et al. (2011): Spatio-temporal transcriptome of the human brain. In: *Nature* 478 (7370), S. 483–489.
- Kanner, L. (1943): Autistic disturbances of affective contact. In: *Nerv. Child* 2, S. 217–250.
- Kirkovski, Melissa; Enticott, Peter G.; Fitzgerald, Paul B. (2013): A Review of the Role of Female Gender in Autism Spectrum Disorders. In: *J Autism Dev Disord.*
- Klauck, Sabine M. (2006): Genetics of autism spectrum disorder. In: *Eur. J. Hum. Genet.* 14 (6), S. 714–720.
- Klein, Thomas; Fischer-Kerli, David (2000): Die Zuverlässigkeit retrospektiv erhobener Lebensverlaufsdaten. Analysen zur Partnerschaftsbiographie des Familiensurveys. In: *Zeitschrift für Soziologie* (4), S. 294–312.
- Knapaska, Ewelina; Kaczmarek, Leszek (2004): A gene for neuronal plasticity in the mammalian brain: *Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK*? In: *Prog. Neurobiol.* 74 (4), S. 183–211.
- Konopka, G.; Wexler, E.; Rosen, E.; Mukamel, Z.; Osborn, G. E.; Chen, L. et al. (2012): Modeling the functional genomics of autism using human neurons. In: *Mol. Psychiatry* 17 (2), S. 202–214.
- Kouzine, Fedor (2007): Supercoil-driven DNA structures regulate genetic transactions. In: *Front Biosci* 12 (8-12), S. 4409.
- Ku, Wei-Chi; Chiu, Sung-Kay; Chen, Yi-Ju; Huang, Hsin-Hung; Wu, Wen-Guey; Chen, Yu-Ju (2009): Complementary quantitative proteomics reveals that transcription factor AP-4 mediates E-box-dependent complex formation for transcriptional repression of *HDM2*. In: *Mol. Cell Proteomics* 8 (9), S. 2034–2050.
- Lai, Cecilia S. L.; Fisher, Simon E.; Hurst, Jane A.; Vargha-Khadem, Faraneh; Monaco, Anthony P. (2001): A forkhead-domain gene is mutated in a severe speech and language disorder. In: *Nature* 413 (6855), S. 519–523.

- Langen, Marieke; Schnack, Hugo G.; Nederveen, Hilde; Bos, Dienke; Lahuis, Bertine E.; Jonge, Maretha V. de et al. (2009): Changes in the developmental trajectories of striatum in autism. In: *Biol. Psychiatry* 66 (4), S. 327–333.
- Le Couteur, A.; Bailey, A.; Goode, S.; Pickles, A.; Robertson, S.; Gottesman, I.; Rutter, M. (1996): A broader phenotype of autism: the clinical spectrum in twins. In: *J Child Psychol Psychiatry* 37 (7), S. 785–801.
- Leclerc, G. M.; Boockfor, F. R.; Faught, W. J.; Frawley, L. S. (2000): Development of a destabilized firefly luciferase enzyme for measurement of gene expression. In: *BioTechniques* 29 (3), S. 590-1, 594-6, 598 passim.
- Lee, Jonathan L. C.; Everitt, Barry J.; Thomas, Kerrie L. (2004): Independent cellular processes for hippocampal memory consolidation and reconsolidation. In: *Science* 304 (5672), S. 839–843.
- Lesch, K. P.; Bengel, D.; Heils, A.; Sabol, S. Z.; Greenberg, B. D.; Petri, S. et al. (1996): Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. In: *Science* 274 (5292), S. 1527–1531.
- Lewandowsky, Stephan; Murdock Jr., Bennet B. (1989): Memory for serial order. In: *Psychological Review* (1), S. 25–57.
- Lichtenstein, Paul; Carlström, Eva; Råstam, Maria; Gillberg, Christopher; Anckarsäter, Henrik (2010): The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. In: *Am J Psychiatry* 167 (11), S. 1357–1363.
- Liu, J.; Nyholt, D. R.; Magnussen, P.; Parano, E.; Pavone, P.; Geschwind, D. et al. (2001): A genomewide screen for autism susceptibility loci. In: *Am. J. Hum. Genet.* 69 (2), S. 327–340.
- Lord, C.; Risi, S.; Lambrecht, L.; Cook, E. H.; Leventhal, B. L.; DiLavore, P. C. et al. (2000): The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. In: *J Autism Dev Disord* 30 (3), S. 205–223.
- Lord, C.; Rutter, M.; Le Couteur, A. (1994): Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. In: *J Autism Dev Disord* 24 (5), S. 659–685.
- Lord, C.; Schopler, E. (1985): Differences in sex ratios in autism as a function of measured intelligence. In: *J Autism Dev Disord* 15 (2), S. 185–193.
- MacDermot, Kay D.; Bonora, Elena; Sykes, Nuala; Coupe, Anne-Marie; Lai, Cecilia S. L.; Vernes, Sonja C. et al. (2005): Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. In: *Am. J. Hum. Genet.* 76 (6), S. 1074–1080.

- Marshall, Christian R.; Noor, Abdul; Vincent, John B.; Lionel, Anath C.; Feuk, Lars; Skaug, Jennifer et al. (2008): Structural variation of chromosomes in autism spectrum disorder. In: *Am. J. Hum. Genet.* 82 (2), S. 477–488.
- Martin, Lisa J.; Gupta, Jayanta; Jyothula, Soma S. S. K.; Butsch Kovacic, Melinda; Biagini Myers, Jocelyn M.; Patterson, Tia L. et al. (2012): Functional Variant in the Autophagy-Related 5 Gene Promotor is Associated with Childhood Asthma. In: *PLoS ONE* 7 (4), S. e33454.
- Martín-Montañez, E.; López-Téllez, J. F.; Acevedo, M. J.; Pavía, J.; Khan, Z. U. (2010): Efficiency of gene transfection reagents in NG108-15, SH-SY5Y and CHO-K1 cell lines. In: *Methods Find Exp Clin Pharmacol* 32 (5), S. 291–297.
- Mülhardt, Cornel (2006): *Der Experimentator: Molekularbiologie, Genomics*. 5. Aufl. München: Elsevier, Spektrum Akad. Verl. (Der Experimentator).
- Mutter, Joachim; Naumann, Johannes; Schneider, Rainer; Walach, Harald; Haley, Boyd (2005): Mercury and autism: accelerating evidence? In: *Neuro Endocrinol. Lett.* 26 (5), S. 439–446.
- Nakabayashi, K.; Scherer, S. W. (2001): The human contactin-associated protein-like 2 gene (CNTNAP2) spans over 2 Mb of DNA at chromosome 7q35. In: *Genomics* 73 (1), S. 108–112.
- Nave, Klaus-Armin (2010): Myelination and support of axonal integrity by glia. In: *Nature* 468 (7321), S. 244–252.
- Newbury, D. F.; Paracchini, S.; Scerri, T. S.; Winchester, L.; Addis, L.; Richardson, Alex J. et al. (2011): Investigation of Dyslexia and SLI Risk Variants in Reading- and Language-Impaired Subjects. In: *Behav Genet* 41 (1), S. 90–104.
- Nord, Alex S.; Roeb, Wendy; Dickel, Diane E.; Walsh, Tom; Kusenda, Mary; O'Connor, Kristen Lewis et al. (2011): Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. In: *Eur. J. Hum. Genet.* 19 (6), S. 727–731.
- Nottebohm, F.; Arnold, A. P. (1976): Sexual dimorphism in vocal control areas of the songbird brain. In: *Science* 194 (4261), S. 211–213.
- Okladnova, O.; Syagailo, Y. V.; Tränitz, M.; Stöber, G.; Riederer, P.; Mössner, R.; Lesch, K. P. (1998): A promoter-associated polymorphic repeat modulates PAX-6 expression in human brain. In: *Biochem. Biophys. Res. Commun.* 248 (2), S. 402–405.
- O'Roak, Brian J.; Deriziotis, Pelagia; Lee, Choli; Vives, Laura; Schwartz, Jerrod J.; Girirajan, Santhosh et al. (2011): Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. In: *Nat Genet* 43 (6), S. 585–589.
- Pardo, Carlos A.; Eberhart, Charles G. (2007): The neurobiology of autism. In: *Brain Pathol.* 17 (4), S. 434–447.

- Peñagarikano, Olga; Abrahams, Brett S.; Herman, Edward I.; Winden, Kellen D.; Gdalyahu, Amos; Dong, Hongmei et al. (2011): Absence of CNTNAP2 Leads to Epilepsy, Neuronal Migration Abnormalities, and Core Autism-Related Deficits. In: *Cell* 147 (1), S. 235–246.
- Peñagarikano, Olga; Geschwind, Daniel H. (2012): What does CNTNAP2 reveal about autism spectrum disorder? In: *Trends in Molecular Medicine* 18 (3), S. 156–163.
- Peter, Beate; Raskind, Wendy H.; Matsushita, Mark; Lisowski, Mark; Vu, Tiffany; Berninger, Virginia W. et al. (2011): Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. In: *J Neurodevelop Disord* 3 (1), S. 39–49.
- Petrin, Aline L.; Giacheti, Célia M.; Maximino, Luciana P.; Abramides, Dagma V. M.; Zanchetta, Sthella; Rossi, Natalia F. et al. (2010): Identification of a microdeletion at the 7q33-q35 disrupting the CNTNAP2 gene in a Brazilian stuttering case. In: *Am. J. Med. Genet.* 152 (12), S. 3164–3172.
- Pickles, A.; Bolton, P.; Macdonald, H.; Bailey, A.; Le Couteur, A.; Sim, C. H.; Rutter, M. (1995): Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. In: *Am. J. Hum. Genet.* 57 (3), S. 717–726.
- Pinto, Dalila; Pagnamenta, Alistair T.; Klei, Lambertus; Anney, Richard; Merico, Daniele; Regan, Regina et al. (2010): Functional impact of global rare copy number variation in autism spectrum disorders. In: *Nature* 466 (7304), S. 368–372.
- Poliak, S. (2003): Juxtaparanodal clustering of Shaker-like K⁺ channels in myelinated axons depends on Caspr2 and TAG-1. In: *The Journal of Cell Biology* 162 (6), S. 1149–1160.
- Poliak, S.; Gollan, L.; Martinez, R.; Custer, A.; Einheber, S.; Salzer, J. L. et al. (1999): Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K⁺ channels. In: *Neuron* 24 (4), S. 1037–1047.
- Poot, Martin; Beyer, Vera; Schwaab, Ira; Damatova, Natalja; Slot, Ruben; Prothero, Jo et al. (2010): Disruption of CNTNAP2 and additional structural genome changes in a boy with speech delay and autism spectrum disorder. In: *Neurogenetics* 11 (1), S. 81–89.
- Pritchard, J. K. (2001): Are rare variants responsible for susceptibility to complex diseases? In: *Am. J. Hum. Genet.* 69 (1), S. 124–137.
- Quandt, K.; Frech, K.; Karas, H.; Wingender, E.; Werner, T. (1995): MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. In: *Nucleic Acids Res.* 23 (23), S. 4878–4884.
- Reichenberg, Abraham; Gross, Raz; Weiser, Mark; Bresnahan, Micheal; Silverman, Jeremy; Harlap, Susan et al. (2006): Advancing paternal age and autism. In: *Arch. Gen. Psychiatry* 63 (9), S. 1026–1032.

- Ronald, Angelica; Happé, Francesca; Bolton, Patrick; Butcher, Lee M.; Price, Thomas S.; Wheelwright, Sally et al. (2006): Genetic heterogeneity between the three components of the autism spectrum: a twin study. In: *J Am Acad Child Adolesc Psychiatry* 45 (6), S. 691–699.
- Rossi, Elena; Verri, Anna Pia; Patricelli, Maria Grazia; Destefani, Valeria; Ricca, Ivana; Vetro, Annalisa et al. (2008): A 12Mb deletion at 7q33–q35 associated with autism spectrum disorders and primary amenorrhea. In: *European Journal of Medical Genetics* 51 (6), S. 631–638.
- Rubenstein, J. L. R.; Merzenich, M. M. (2003): Model of autism: increased ratio of excitation/inhibition in key neural systems. In: *Genes Brain Behav.* 2 (5), S. 255–267.
- Rubenstein, John L. R. (2011): Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. In: *J Child Psychol Psychiatry* 52 (4), S. 339–355.
- Sanger, F.; Nicklen, S.; Coulson, A. R. (1992): DNA sequencing with chain-terminating inhibitors. 1977. In: *Biotechnology* 24, S. 104–108.
- Sasson, Noah J.; Lam, Kristen SI; Parlier, Morgan; Daniels, Julie L.; Piven, Joseph (2013): Autism and the broad autism phenotype: familial patterns and intergenerational transmission. In: *J Neurodev Disord* 5 (1), S. 11.
- Sawaya, Sterling M.; Bagshaw, Andrew T.; Buschiazzo, Emmanuel; Gemmell, Neil J. (2012): Promoter microsatellites as modulators of human gene expression. In: *Adv. Exp. Med. Biol.* 769, S. 41–54.
- Sawaya, Sterling; Bagshaw, Andrew; Buschiazzo, Emmanuel; Kumar, Pankaj; Chowdhury, Shantanu; Black, Michael A. et al. (2013): Microsatellite Tandem Repeats Are Abundant in Human Promoters and Are Associated with Regulatory Elements. In: *PLoS ONE* 8 (2), S. e54710.
- Schopler, E.; Andrews, C. E.; Strupp, K. (1979): Do autistic children come from upper-middle-class parents? In: *J Autism Dev Disord* 9 (2), S. 139–152.
- Scott-Van Zeeland, A. A.; Abrahams, B. S.; Alvarez-Retuerto, A. I.; Sonnenblick, L. I.; Rudie, J. D.; Ghahremani, D. et al. (2010): Altered Functional Connectivity in Frontal Lobe Circuits Is Associated with Variation in the Autism Risk Gene CNTNAP2. In: *Science Translational Medicine* 2 (56), S. 56ra80.
- Sebat, Jonathan; Lakshmi, B.; Malhotra, Dheeraj; Troge, Jennifer; Lese-Martin, Christa; Walsh, Tom et al. (2007): Strong association of de novo copy number mutations with autism. In: *Science* 316 (5823), S. 445–449.

- Sehested, Line T.; Møller, Rikke S.; Bache, Iben; Andersen, Noemi B.; Ullmann, Reinhard; Tommerup, Niels; Tümer, Zeynep (2010): Deletion of 7q34-q36.2 in two siblings with mental retardation, language delay, primary amenorrhea, and dysmorphic features. In: *Am. J. Med. Genet.* 152 (12), S. 3115–3119.
- Shao, Yujun; Wolpert, Chantelle M.; Raiford, Kimberly L.; Menold, Marisa M.; Donnelly, Shannon L.; Ravan, Sarah A. et al. (2002): Genomic screen and follow-up analysis for autistic disorder. In: *Am. J. Med. Genet.* 114 (1), S. 99–105.
- Shastry, Barkur S. (2009): SNPs: impact on gene function and phenotype. In: *Methods Mol. Biol.* 578, S. 3–22.
- Stanfield, Andrew C.; McIntosh, Andrew M.; Spencer, Michael D.; Philip, Ruth; Gaur, Sonia; Lawrie, Stephen M. (2008): Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies. In: *Eur. Psychiatry* 23 (4), S. 289–299.
- Steffenburg, S.; Gillberg, C.; Hellgren, L.; Andersson, L.; Gillberg, I. C.; Jakobsson, G.; Bohman, M. (1989): A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. In: *J Child Psychol Psychiatry* 30 (3), S. 405–416.
- Strauss, Kevin A.; Puffenberger, Erik G.; Huentelman, Matthew J.; Gottlieb, Steven; Dobrin, Seth E.; Parod, Jennifer M. et al. (2006): Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. In: *N. Engl. J. Med.* 354 (13), S. 1370–1377.
- Strick, Peter L.; Dum, Richard P.; Fiez, Julie A. (2009): Cerebellum and nonmotor function. In: *Annu. Rev. Neurosci.* 32, S. 413–434.
- Szatmari, Peter; Paterson, Andrew D.; Zwaigenbaum, Lonnie; Roberts, Wendy; Brian, Jessica; Liu, Xiao-Qing et al. (2007): Mapping autism risk loci using genetic linkage and chromosomal rearrangements. In: *Nat Genet* 39 (3), S. 319–328.
- Tan, Geoffrey C. Y.; Doke, Thomas F.; Ashburner, John; Wood, Nicholas W.; Frackowiak, Richard S. J. (2010): Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of CNTNAP2. In: *Neuroimage* 53 (3), S. 1030–1042.
- Teramitsu, I. (2006): FoxP2 Regulation during Undirected Singing in Adult Songbirds. In: *Journal of Neuroscience* 26 (28), S. 7390–7394.
- Teramitsu, Ikuko; Kudo, Lili C.; London, Sarah E.; Geschwind, Daniel H.; White, Stephanie A. (2004): Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. In: *J. Neurosci.* 24 (13), S. 3152–3163.
- Teramitsu, Ikuko; Poopatanapong, Amy; Torrisi, Salvatore; White, Stephanie A.; Tanimoto, Hiromu (2010): Striatal FoxP2 Is Actively Regulated during Songbird Sensorimotor Learning. In: *PLoS ONE* 5 (1), S. e8548.

- Thompson, J. F.; Hayes, L. S.; Lloyd, D. B. (1991): Modulation of firefly luciferase stability and impact on studies of gene regulation. In: *Gene* 103 (2), S. 171–177.
- Tuchman, Roberto; Rapin, Isabelle (2002): Epilepsy in autism. In: *Lancet Neurol* 1 (6), S. 352–358.
- Uhlhaas, Peter J.; Singer, Wolf (2006): Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. In: *Neuron* 52 (1), S. 155–168.
- Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G. et al. (2001): The sequence of the human genome. In: *Science* 291 (5507), S. 1304–1351.
- Vernes, Sonja C.; Newbury, Dianne F.; Abrahams, Brett S.; Winchester, Laura; Nicod, Jérôme; Groszer, Matthias et al. (2008): A Functional Genetic Link between Distinct Developmental Language Disorders. In: *N Engl J Med* 359 (22), S. 2337–2345.
- Volkmar, Fred R.; Lord, Catherine; Bailey, Anthony; Schultz, Robert T.; Klin, Ami (2004): Autism and pervasive developmental disorders. In: *J Child Psychol Psychiatry* 45 (1), S. 135–170.
- Wechsler, David (1949): Manual for the Wechsler Intelligence Scale for Children. Psychological Corporation. New York.
- Weintraub, Karen (2011): The prevalence puzzle: Autism counts. In: *Nature* 479 (7371), S. 22–24.
- Whitehouse, A. J. O.; Bishop, D. V. M.; Ang, Q. W.; Pennell, C. E.; Fisher, S. E. (2011): CNTNAP2 variants affect early language development in the general population. In: *Genes, Brain and Behavior* 10 (4), S. 451–456.
- Williams, Heather (2004): Birdsong and singing behavior. In: *Ann. N. Y. Acad. Sci.* 1016, S. 1–30.
- Windham, Gayle C.; Zhang, Lixia; Gunier, Robert; Croen, Lisa A.; Grether, Judith K. (2006): Autism spectrum disorders in relation to distribution of hazardous air pollutants in the san francisco bay area. In: *Environ. Health Perspect.* 114 (9), S. 1438–1444.
- Wing, Lorna (1981): Asperger syndrome: A clinical account. In: *Psychol Med* (11), S. 115–129.
- Wing, Lorna; Potter, David (2002): The epidemiology of autistic spectrum disorders: is the prevalence rising? In: *Ment Retard Dev Disabil Res Rev* 8 (3), S. 151–161.
- World Health Organization (1992): International classification of mental and behavioral disorders. Clinical descriptions and diagnostic guidelines. World Health Organization. Geneva.
- Zhang, Gong; Hubalewska, Magdalena; Ignatova, Zoya (2009): Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. In: *Nat Struct Mol Biol* 16 (3), S. 274–280.

10 Anhang

10.1 In dieser Studie verwendete Datenbanken

dbSNP: Version 3.1, March 8, 2006; <http://www.ncbi.nlm.nih.gov/SNP/>

ENCODE: June 2005 Freeze; <http://genome.ucsc.edu/ENCODE/>

Ensemble: Ensembl release 56 – September 2009; <http://www.ensembl.org/index.html>

Hapmap: Data Rel 27 Phase+III; <http://www.hapmap.org>

MatInspector: genomatrix software suite v2.6; <http://www.genomatix.de>

USCS: Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly; <http://genome.ucsc.edu/>

Watcut: June 19, 2007. Enzymes as of July 2006, Rebase; <http://watcut.uwaterloo.ca/>

10.2 Abkürzungsverzeichnis

%(v)	Volumenprozent
%(w)	Massenprozent
µg	Mikrogramm
µl	Mikroliter
µM	Mikromol
A	Autismus
AG	Aktiengesellschaft
AGRE	Autism Genetic Resource Exchange
Amp	Ampicillin
APS	Ammoniumperoxidsulfat
Aqua ad iniectionabilia	Wasser für Injektionszwecke
Aquabidest	bidestilliertes Wasser
ASS	Autismus-Spektrum-Störung
ATCC	American Type Culture Collection
bp	Basenpaare
BSA	Bovines Serumalbumin
bzw.	beziehungsweise
C	Konzentration
cAMP	cyclisches Adenosinmonophosphat
Chi ²	Chi ² -Verteilung
CI	Konfidenzintervall

CNTNAP2	Contactin-associated protein-like 2
CPBP	Core promoter-binding protein
CTCF	Insulator protein CTCF (Bindefaktor für CCCTC)
ddNTP	Didesoxyribonukleosidtriphosphat
DKFZ	Deutsches Krebsforschungszentrum
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamid
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleosidtriphosphat
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders
E.coli	Escherichia Coli
E230	Extinktion bei 230nm
E260	Extinktion bei 260nm
E280	Extinktion bei 280nm
EDTA	Ethylendiamintetraacetat
EGR1	Early growth response 1
ER	Estrogen response element
F	Verdünnungsfaktor
FAUT	Autismusproben Frankfurt
FBS	Fetal Bovine Serum
FCD	fokale corticale Dysplasie
FOXP2	Forkhead box P2
GFP	Grün Fluoreszierendes Protein
GLI3	GLI-Kruppel family member 3
GmbH	Gesellschaft mit beschränkter Haftung
GRCh37	Genome Reference Consortium human genome (build 37)
GWAS	Genomweite Assoziationsstudie
H ₂ O	Dihydrogenmonoxid (Wasser)
HAUT	Autismusproben Homburg
hBDNF	Brain-derived neurotrophic factor human
HFA	High-Functioning-Autismus
HMFM	Hogness Modified Freezing Medium
ICD-10	International Statistical Classification of Diseases and Related Health Problems

IQ	Intelligenzquotient
KCl	Kaliumchlorid
kD	Kilodalton
KGaA	Kommanditgesellschaft auf Aktien
KJPPP	Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes und Jugendalters
KO	Knockout
KRAB	Krüppel associated box
LB	Lysogeny Broth
LE	Low Electroendosmosis
LFA	Low-Functioning-Autismus
LOD	Logarithm of the odds
m	männlich
M	Mol
MAF	Minorallelfrequenz
Mb	Megabasen
MEM	Minimal Essential Medium
MgCl ₂	Magnesiumchlorid
Min	Minute
ml	Milliliter
mM	Millimol
mPFC	medialer präfrontaler Cortex
MRT	Magnetresonanztomographie
MYOG	Myogenic bHLH protein myogenin (myf4)
NaCl	Natriumchlorid
NCBI	National Center of Biotechnology
Ng	Nanogramm
NGFIC	Nerve growth factor-induced protein C
Nm	Nanometer
NN	Nicht nachweisbar
NRSF	Neuron-restrictive silencer factor
NV	Nicht vorhanden
OH	Hydroxylgruppe
OR	Odds Ratio
p.a.	pro analysi (Reinheitsgrad für Chemikalien)

PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerasekettenreaktion
PDD-NOS	Pervasive developmental disorder not otherwise specified
pH	potential Hydrogenii
pIZ	nach Induktion der Zelldifferenzierung
RFLP	Restriktionsfragmentlängenpolymorphismus
Rpm	revolutions per minute
S	Spektrumstörung
SD	Standardabweichung
SDS	Sodiumlaurylsulfat
SNP	Einzelnukleotidpolymorphismus
SRD	Scientific Research and Development GmbH (Bad Homburg)
SSES	Spezifische Sprachentwicklungsstörung
STR	Short Tandem Repeat
Taq	Thermus aquaticus
TDT	Transmission Disequilibrium Test
TEMED	Tetramethylethylenamin
TE-Puffer	Tris-EDTA Puffer
TF	Transkriptionsfaktor
TFAP4	Activator protein 4
TFBS	Transkriptionsfaktorbindestellen
Tris	Tris(hydroxymethyl)-aminomethan
TSS	Transcription start site
TTE-Puffer	TRIS-Taurin-EDTA Puffer
U-test	Mann-Whitney-U-Test
UV	Ultraviolett
w	weiblich
X-Gal	5-Brom-4-chlor-3-indoxyl- β -D-galactopyranosid
ZF219	Kruppel-like zinc finger protein 219
ZF300	KRAB-containing zinc finger protein 300
ZF9	Core promoter-binding protein (CPBP) mit 3 Kruppel-type Zinkfinger (auch bekannt als KLF6)

10.3 Sequenz der *CNTNAP2*-Promotorregion

Homo sapiens chromosome 7 genomic contig, GRCh37 reference primary assembly

NCBI Reference Sequence: NT_007914.15

>ref|NT_007914.15|:6407832-6409832 Homo sapiens chromosome 7 genomic contig, GRCh37 reference primary assembly

```

AAAACTCAGTTTATTCCAGGATTAGGCAAGCCTCCATAAATTTTCTATTTATCTTTAAATGCAAAT
GCTTATTTCATTGCATAACATCGAAAGACATGGTTACTCCTACTTAGTCCTCCACAAACATCATTAG
ATATTTAGATGACCAATAATTTTAAATTGTGTCAATAAAGTCAGTGGAATGCTTGGCTCTGCAGAG
ACATAAGAATTCAAATACTTTAAGTAGACTAGATGGAAGAAAAACATATGGAGGGTAGAGTGTGAA
GAGCTGTGAGCTGGGGAGAAGAGCTTGATAATTATTATTAATAATTATCTAAATGTGAAAAAATCTA
GGTCTCTCTAAACATTCTTAATTCTTAGTGCTTGCAATGCCTCAGGTACCTTGAAATAGTTACCCA
TATGCCTGCTTATGTATTTCCGTTTTGTGTTTGTGTTTGTATGCGTGTTTGCCTAAGTGTAGTTTC
TCCTGGCAGCAGCGTGTGTTTGGTGTACGTGTGCATATGTATGAGTACGAATTAATTACATATCGA
AAACATGTGCATTAACGGCATCTGCAGTCCTCCCTCCCTGCTCCAGTCAGTCGGGATGTTTAGCAC
CGTCACCTTGGTGTTAGGACCACTCGGAAGACCTGCCTGTCGGACGCTGACTGTGGTTCCCGGGGG
GCCGGAGCTCCAGGAGCCTGTACCATCTTTATCTGAAGGGGAGGAGAGAAGGAAAGGAGAAGGCAT
CCGAGCGAGGGTGTTCCTTCCTGCAGCCGGGTAGGGGGCAGCCAGGTTTGGAGCAGGGTTCTAG
GATTACACGGCAAATCCCCAATTTCAGCACCGCGGACAGCATCACCTCCCGCCCAGGGAGAGGCGC
GCGCCTCCTTGCTCCGCCCCCGCCCCAGCTGATGGACCCCGGCCGGTCGGCTCAGTCTCCCTTTTG
CTTTCAAACGAGGCCAGGGCAGGGCAGAGGGTTTTGACAGCCAGAGGGGTACGGTAAGGAGGAGG
AAGAGGAGGAGGAGGCGGAGGAGGAGGAGGAGCCGAGGAGGGAACCGAGAAGGGAAGAGAAGGGAG
GGCTTCCCTCCGCTCGCTCGCCGCCCGGGCTCACAGGAGCTGGCGGCGGCGGCGGCGGCGGCGG
GTCTCCTGCTCTCCGCGGCTGTTGCTGCTGCTCGCGCTCCGCCGCCCGGGAGATGCTTCCTCGCGC
GGCGCAGCGCTGAGGCCGTGCGTGCGCCCCGGCTGCGCTGCGCGCTCCCCACATACACAAGCTCTC
CATGTGAGCTGACAGGCGAGTGGAACCCCTCGAGTCACGCTGCCCGGCGGCGGAGGGAGCGCTCG
CCCGCAGTGGAACAGCTGCACCACCGTCCCCGTGCTCTGCCTTCCTCTTCTGCAGCCTCTGCTC
TTCTGATTACCTCCCTCCCCCGTCTTTGGTGATTTTTTTTTTTTCAAGAAGGAGAGGGCGGGGTAG
GTGTCCGTTCCCTCCCCTCTTCCCCCTCCTTTGCCTTCTTGGTTTGAATTTCTCCCCCGGCGTTG
CACTGGCACACAGTGCAAGAGGCAATACCCGCACGGAGGGAGAACGAAGGCTGAGACTCCCCTGCC
GCTCCAAGCCCGGAAGAACTGGAGCCTGGAGGGGGGTGAGGGGAGAAGAGGAAGCGGGAGGGGGCTT
GGCTTCCTCGCGTATTTGAGGACAGCCCATCTCCCTTCAAGAACCCTACGGAGAGTCGGAATGCAT
CTCCGCAGCGAGCTCTTGGAGCGCCGCCGGCGGGAGGCGAAGGATGCAGGCGGCTCCGCGCGCCC
GCTGCGGGGCAGCGCTCCTGCTGTGATTGTGAGCAGCTGCCTCTGCAGAGCCTGGACGGCTCCCT
CCACGTCCCGTAAGTAGCCGTCTCCTCGCTCTGCTCTGGAGCAGTTTCAGTGCGGCATTGATGTTT
GGCACCAGGGTATCAACTCCGAAGTGCATCGCAGTGCTGGCACCCCTGATGTGTTTGTGTCTCCGCT
GTCACACACTTGCAGCCACTG

```

11 Danksagung

Da ich befürchte, dass ich an dieser Stelle womöglich irgendjemanden unbeabsichtigt vergesse, möchte ich zunächst erst einmal *ALLEN* danken, die mich in den letzten Jahren unterstützt haben und somit maßgeblich an der Entstehung dieser Arbeit beteiligt waren.

Das größte Dankeschön gilt jedoch meinen Eltern *KARIN* und *HARRY*, sowie meiner Schwester *SANDRA*, die mich nicht nur in den letzten Jahren, sondern während meines ganzen Lebens mit offenen Ohren, guten Tipps und einer großen Portion Unterstützung in allen Lebenslagen begleitet haben. Ohne sie wäre nicht nur das letzte Jahr sehr schwer für mich gewesen. Nicht besonders leicht war diese Zeit wohl auch für *ARNE*, der mich in meiner kreativen Phase des Schreibens trotz meiner Launen und Stimmungsschwankungen ertragen hat und mich immer wieder ermutigte und motivierte, wofür ich ihm ebenfalls von Herzen danke. Meinen beiden Nichten *VIVIAN* und *AMÉLIE* danke ich dafür, dass sie auch in besonders stressigen Zeiten immer dafür gesorgt haben, dass ich pflichtbewusst meine Pausen einhalte und dass sie mir in dieser Zeit stets ein Lächeln auf die Lippen gezaubert haben.

Bei Frau *CHRISTINE M. FREITAG* bedanke ich mich dafür, dass sie mir die Möglichkeit gegeben hat, bei dem Aufbau des Molekulargenetischen Labors der Kinder- und Jugendpsychiatrie entscheidend mitzuwirken und das ich meine Doktorarbeit an ihrem Institut durchführen konnte. Vielen Dank auch für die unterschiedlichen Fortbildungen die sie mir ermöglicht haben, sowie für all die guten Gespräche, Diskussionen und Ratschläge in den vergangenen 4 ½ Jahren.

Ebenso danke ich *BODO LAUBE* und *RALF GALUSKE*, die sich sofort dazu bereit erklärt hatten, die Betreuung und Gutachtenerstellung meiner Arbeit zu übernehmen.

ANDREAS GEBURTIQ-CHIOCCHETTI möchte ich zum einen für seine unglaubliche Geduld, die er während seiner nunmehr zweijährigen Mitbetreuung meiner Arbeit mit mir hatte, danken. In dieser Zeit hat er nämlich nicht nur mir durch sein fast schon autismustypisches repetitives Verhalten die Grundlagen der Statistik und der damit verbundenen Anwendungsprogramme näher gebracht. Zum

anderen möchte ich ihm für die sehr penible Einarbeitung in die Grundlagen der Zellkultur sowie für die mehrfache Korrektur aller von mir angefertigter Schriftstücke ganz herzlich danken.

Ein großes Dankeschön geht auch an *ALLE FAMILIEN*, die sich dazu bereit erklärt haben, uns ihre Erbinformationen zur Verfügung zu stellen, sowie an *ALLE KLINIKER*, die am Sammeln und Bereitstellen der Proben und der dazugehörigen Daten beteiligt waren. Vielen Dank auch *HEIKO ZERLAUT* für die Zusammenstellung der Phänotypdaten. Ohne euch wäre die Durchführung dieser Arbeit gar nicht erst möglich gewesen.

SABINE M. KLAUCK möchte ich für die gute Zusammenarbeit danken und *FRAU WIRTH* für die rasche Durchführung der Analysen an den damals noch in Heidelberg gelagerten Proben.

Bei Frau *SIMONE FULDA* und *ULRIKE GRAAB* möchte ich für die Bereitstellung der humanen Zelllinien und für die Einweisung in die Methodik des Luziferase-Assays bedanken.

Ein großes Dankeschön geht auch an *DENISE HASLINGER*, deren exzellente Ergebnisse ich zu einem kleinen Teil im Rahmen meiner Arbeit vorstellen und als Diskussionsgrundlage nutzen durfte.

Zum Schluss möchte ich noch *SILVIA* und *REGINA* aus dem „Wohlfühlabor“ für die schöne Zeit im Feenzimmer danken. Nicht nur das immer währende Vorhandensein an Süßwaren, sondern auch die Freundschaft zu euch hat mir die Arbeit an meinem Projekt deutlich erleichtert. Oder waren es vielleicht doch unsere konstruktiven Gespräche sowie euer fundiertes Fachwissen, die mir bei der Lösung meiner Probleme weitergeholfen haben? Was auch immer...zumindest danke ich euch für alles!!

12 Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

A handwritten signature in black ink, appearing to read 'M. Kopp', with a stylized flourish at the end.

Darmstadt, den 28.10.2013

13 Lebenslauf



Marnie Kopp

Von-der-Au-Straße 20
64297 Darmstadt

Geboren in Groß-Gerau am 19.04.1978

Akademische Ausbildung

April 2009 – Juni 2013

Promotion; Molekulargenetisches Labor der Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes- und Jugendalters, Klinikum der Johann Wolfgang Goethe-Universität, Frankfurt am Main

Mai 2007 – Februar 2008

Diplomarbeit; Zentrum der Rechtsmedizin; Klinikum der Johann Wolfgang Goethe-Universität; Frankfurt am Main

Oktober 2000 – Februar 2008

Studium zur Diplombiologin; Technische Universität Darmstadt; Darmstadt

Arbeitserfahrung

April 2009 – Juni 2013

Wissenschaftliche Mitarbeiterin; Molekulargenetisches Labor der Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes- und Jugendalters, Klinikum der Johann Wolfgang Goethe-Universität, Frankfurt am Main

März 2008 – März 2009

Beschäftigung im DNA-Labor zur Unterstützung der Routine und der Forschung; Zentrum der Rechtsmedizin; Klinikum der Johann Wolfgang Goethe-Universität; Frankfurt am Main

Juli 2007 – Februar 2008

Studentische Aushilfskraft im Routinelabor; Zentrum der Rechtsmedizin; Klinikum der Johann Wolfgang Goethe-Universität; Frankfurt am Main

Berufsausbildung

Oktober 1997 – September 2000

Ausbildung zur Medizinisch-Technischen-Laboratoriumsassistentin; Schule für Technische Assistenten in der Medizin; Städtische Kliniken; Frankfurt am Main – Höchst

Weiterbildung

September 2010

Fortbildungsveranstaltung für Projektleiter/innen und Beauftragte für die biologische Sicherheit gentechnischer Anlagen; Sicherheit in der Gentechnik; Johannes Gutenberg-Universität; Mainz

Fachkenntnisse

Methodenspektrum

Auflichtmikroskopie, DHPLC, Durchlichtmikroskopie, DNA-Extraktion, Gelelektrophorese (horizontal und vertikal), Fluoreszenzmikroskopie, Herstellung histologischer Präparate, Kapillarsequenzierung und Auswertung der Sequenzierdaten, Klonierung, Luziferase-Assay, PCR, Real-Time PCR, RFLP, TRFLP, Western Blot, Zellkultur

EDV

GeneScan Analysis 3.7 (ABI), Genotyper 3.6 NT (ABI), Haploview, ImageJ, Microsoft Office, Photoshop CS, Primer Express, R, SeqScape 2.5 (ABI), Sequencing Analysis 2.1.2 (ABI), Sequence Navigator 1.0.1 (ABI), SPSS

P u b l i k a t i o n e n

Kauferstein S, Kiehne N, Jenewein T, Biel S, Kopp M, König R, Erkapic D, Rothschild M, Neumann T (2013) **Genetic analysis of sudden unexplained death: A multidisciplinary approach.** Forensic Sci Int Genet. 2013 Jun; 299(1-3):122-7

Leblond CS, Heinrich J, Delorme R, Proepper C, Betancur C, Huguet G, Konyukh M, Chaste P, Ey E, Rastam M, Anckarsäter H, Nygren G, Gillberg IC, Melke J, Toro R, Regnault B, Fauchereau F, Mercati O, Lemièrre N, Skuse D, Poot M, Holt R, Monaco AP, Järvelä I, Kantojärvi K, Vanhala R, Curran S, Collier DA, Bolton P, Chiocchetti A, Klauck SM, Poustka F, Freitag CM, Waltes R, Kopp M, Duketis E, Bacchelli E, Minopoli F, Ruta L, Battaglia A, Mazzone L, Maestrini E, Sequeira AF, Oliveira B, Vicente A, Oliveira G, Pinto D, Scherer SW, Zelenika D, Delepine M, Lathrop M, Bonneau D, Guinchat V, Devillard F, Assouline B, Mouren MC, Leboyer M, Gillberg C, Boeckers TM, Bourgeron T (2012) **Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders.** PLoS Genet. 2012 Feb;8(2):e1002521



Detailed Status Information

Manuscript #	2013MP000531
Current Revision #	0
Submission Date	18th Jun 13
Current Stage	
Title	Variants of the CNTNAP2 promoter as risk factors for autism spectrum disorders: a genetic and functional approach
Running Title	CNTNAP2 promoter variants in ASD
Manuscript Type	Original Article
Manuscript Comment	<p>We, including all co-authors, are accepting the condition that Molecular Psychiatry levies page charges should the paper be accepted and published.</p> <p>Please note that the submission system has automatically and incorrectly changed the affiliations of two authors (i.e. AG Chiocchetti, D Haslinger). Furthermore, the affiliations of A Voran, U Graab and S Fulda were not correctly displayed in the validation setting, although being correct in the author details section. For correct affiliations please refer to the title page in the main manuscript.</p>
Corresponding Author	Prof. Christine Freitag (University Hospital Frankfurt)
Contributing Authors	Mrs. Marnie Kopp , Dr. Andreas G. Chiocchetti , Dr. Regina Waltes , Mrs. Denise Haslinger , Dr. Eftichia Duketis , Dr. Tomasz Jarzok , Prof. Fritz Poustka , Dr. Anette Voran , Mrs. Ulrike Graab , Prof. Jobst Meyer , Dr. Sabine M. Klauck , Prof. Simone Fulda
Abstract	<p>CNTNAP2 (contactin-associated protein-like 2 gene), a member of the Neurexin gene superfamily, is one of the best-replicated risk genes for autism spectrum disorders (ASD). ASD are predominately genetically determined neurodevelopmental disorders characterized by impairments of language development, social interaction and communication, as well as stereotyped behavior and interests. Although CNTNAP2 expression levels were proposed to alter ASD risk, no study to date has focused on its promoter. Here, we directly sequenced the CNTNAP2 promoter region of 667 members of 236 German families with one child diagnosed with ASD. We genotyped the three most frequent variants (rs150447075, rs34712024, rs71781329) in additional 356 families (1102 members). Functional analyses of these variants were performed in HEK293T cells and during neuronal differentiation of SH-SY5Y cells. The minor G allele of rs34712024 was under-transmitted in ASD subjects and showed a trend for association with language development. Carriers of rs71781329GCG[7] showed a delayed onset of speech. At the functional level, all variants were predicted to alter transcription factor binding sites. In a luciferase assay, the minor alleles showed cell line-specific and differentiation stage-dependent effects at the level of promoter activation. Interestingly, only variants that altered transcription factor binding sites related to neuronal activity (i.e. rs34712024 and rs71781329) showed an effect on ASD liability or language development. We conclude that the effect of CNTNAP2 promoter variants on ASD risk and language development is pleiotropic. This pleiotropy may be mediated by the specific cellular context and differential transcription factor activation during neuronal differentiation.</p>
Keywords	Association study, 7q35-q36, transcriptional efficiency, protective effect, language development
Conflict of Interest Statement	<p>The authors have declared there is NO conflict of interest to disclose</p> <p>None of the authors has to report any conflict of interest. Talks of Christine Freitag have been supported by Shire, Novartis, and Eli Lilly; and she has acted as consultant for Desitin.</p>

Stage	Start Date
Reviewer Assignment	18th Jun 13
In Submission	18th Jun 13
Pre Submission	18th Jun 13

Variants of the *CNTNAP2* promoter as risk factors for autism spectrum disorders: a genetic and functional approach

Running title: *CNTNAP2* promoter variants in ASD

Marnie Kopp, MSc¹, Andreas Chiocchetti, PhD¹, Regina Waltes, PhD¹, Denise Haslinger, MSc¹, Eftichia Duketis, MD¹, Tomasz Jarczok, MD¹, Fritz Poustka, MD¹, Anette Voran, MD², Ulrike Graab, MSc³, Jobst Meyer, PhD⁴, Sabine M. Klauck, PhD⁵, Simone Fulda, MD³, Christine M. Freitag, MD^{1*}

¹ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, JW Goethe University, Frankfurt am Main, Germany

² Department of Child and Adolescent Psychiatry, Saarland University, Homburg, Germany

³ Institute of Experimental Cancer Research in Pediatrics, Frankfurt am Main, Germany

⁴ Department of Neurobehavioral Genetics, Institute of Psychobiology, University of Trier, Germany

⁵ Division of Molecular Genome Analysis, German Cancer Research Center (DKFZ), Heidelberg, Germany

* Corresponding author

Professor Dr. med. Dipl. theol. Christine M. Freitag, MD, MA

Klinik für Psychiatrie, Psychosomatik und Psychotherapie des Kindes und Jugendalters

Deutschordenstr. 50

60528 Frankfurt am Main

Germany

T: +49 – 69 - 6301-5408

F: +49 – 69 - 6301-5843

e-mail: c.freitag@em.uni-frankfurt.de

Grants:

The study was in part supported by grant Po 255/17-4 of the Deutsche Forschungsgemeinschaft to F. Poustka and ERA-NET NEURON BMBF grant EUHFAUTISM-01EW1105 to C.M. Freitag.

Abstract

CNTNAP2 (contactin-associated protein-like 2 gene), a member of the Neurexin gene superfamily, is one of the best-replicated risk genes for autism spectrum disorders (ASD). ASD are predominately genetically determined neurodevelopmental disorders characterized by impairments of language development, social interaction and communication, as well as stereotyped behavior and interests. Although *CNTNAP2* expression levels were proposed to alter ASD risk, no study to date has focused on its promoter. Here, we directly sequenced the *CNTNAP2* promoter region of 667 members of 236 German families with one child diagnosed with ASD. We genotyped the three most frequent variants (rs150447075, rs34712024, rs71781329) in additional 356 families (1102 members). Functional analyses of these variants were performed in HEK293T cells and during neuronal differentiation of SH-SY5Y cells. The minor G allele of rs34712024 was under-transmitted in ASD subjects and showed a trend for association with language development. Carriers of rs71781329GCG[7] showed a delayed onset of speech. At the functional level, all variants were predicted to alter transcription factor binding sites. In a luciferase assay, the minor alleles showed cell line-specific and differentiation stage-dependent effects at the level of promoter activation. Interestingly, only variants that altered transcription factor binding sites related to neuronal activity (i.e. rs34712024 and rs71781329) showed an effect on ASD liability or language development. We conclude that the effect of *CNTNAP2* promoter variants on ASD risk and language development is pleiotropic. This pleiotropy may be mediated by the specific cellular context and differential transcription factor activation during neuronal differentiation.

Keywords

Association study; 7q35-q36; transcriptional efficiency; protective effect; language development

Introduction

Autism spectrum disorders (ASD) are childhood onset neurodevelopmental disorders, showing a prevalence of approximately 1% in the general population.^{1,2} ASD are characterized by pervasive impairments of social interaction and communication, repetitive and stereotyped patterns of behavior and restricted interests.^{3,4} They comprise childhood autism, Asperger's Syndrome and atypical autism respective pervasive developmental disorder - not otherwise specified (PDD-NOS). Although twin studies have shown a heritability of 70~90%^{5,6} the genetic architecture of ASD has not yet been fully elucidated.

The contactin-associated protein-like 2 gene (*CNTNAP2*), a member of the neurexin gene superfamily, has been suggested as one of the most promising risk genes for ASD and language development.⁷ It is located within one of the best replicated chromosomal regions detected by linkage studies in ASD (7q22-q36)⁶ and language development.⁸ In addition, several studies on *CNTNAP2* have reported association of genetic variants with ASD and language development. Increased risk for ASD for rs7794745T was noted in a study combining linkage- and subsequent family-based association methods. Furthermore, this variant of intron 2 of *CNTNAP2* displayed a male-specific and a parent-of-origin effect,⁹ partly explaining the marked sex differences observed in the incidence of autism. In addition, the major C allele of the non-coding variant rs2710102 (intron 13) increased ASD risk. Again, a mainly male-driven association with age at first spoken word was observed.¹⁰ The relevance of *CNTNAP2* variants for language development is further highlighted by an association of rs2710102C with poor non-word repetition test performance and lower receptive and expressive language abilities in specific language impairment (SLI).^{11,12} In an Australian general population sample, a four-marker haplotype including rs2710102C correlated with language acquisition. The homozygous carriers of the haplotype obtained substantially lower scores in the communication subscale of the Infant Monitoring Questionnaire.¹³ Functionally, a magnetic resonance imaging (fMRI) study showed that healthy individuals homozygous for both ASD risk alleles rs7794745T and rs2710102C exhibited an increased activation in the right inferior frontal gyrus during a verbal fluency

task, which typically induces activation of the left side, i.e. Broca's area.¹⁴ This atypical language lateralization has also been observed in ASD.¹⁵

Sequencing studies of *CNTNAP2* found an elevated burden of rare and private mutations in patients with ASD compared to controls.^{16,17} Most of the identified deleterious or non-synonymous mutations resided in the C-terminal part of *CNTNAP2* with no specific enrichment in any single domain (Exons 14-24). Their functional impact is still elusive. Also, several case studies described rare deletions of *CNTNAP2*. For most of the patients language related deficits were reported, including stuttering, delay of language or ASD.¹⁸⁻²² One individual with ASD carried a maternally inherited rare deleterious CNV within the promoter region of *CNTNAP2*, which lead to reduced *CNTNAP2* mRNA expression in lymphoblastoid cell lines of carriers with ASD in this family.²³

CNTNAP2 knock out mice showed behavioral phenotypes reminiscent of ASD including stereotypic motor movements, behavioral inflexibility, abnormalities in social behavior, and reduced ultrasonic vocalization. At the cellular level of the brain, an aberrant neuronal migration, reduction of GABAergic interneurons, and aberrant neuronal synchrony were described, pointing towards a central role of *CNTNAP2* in early brain development and neuronal differentiation.²⁴ The central role of this protein is further highlighted by its role in the clustering of Potassium (K⁺)-channels at the juxtaparanodal region of the nodes of Ranvier²⁵, and its high expression levels in language related brain areas.¹⁰

Taken together, there is convergent evidence from linkage, association, CNV, knock-out and expression studies for involvement of *CNTNAP2* in ASD and language related phenotypes. It is likely that genetic variants or a reduced gene dosage of this susceptibility gene modulate ASD and/or language development by changing *CNTNAP2* functionality and availability during neuronal development. Despite the obvious functional impact of promoter variants on mRNA expression levels,^{26,27} no study to date has focused on the functional role of *CNTNAP2* promoter variants in the context of ASD.

Therefore, we aimed at testing variants of the *CNTNAP2* promoter for association with ASD and language development in a large parent-child sample and characterizing their effects on transcriptional activity during neuronal differentiation in SH-SY5Y cell lines and mitotically active non-neuronal HEK293T cells. Here, we report a nominal association of the *CNTNAP2* promoter variant rs34712024 with ASD. Furthermore, two of the analyzed variants may also influence language development. Characterization of the transcriptional efficiency of the promoter revealed a specific activation during late stages of neuronal differentiation and cell type-specific effects of all investigated variants.

Materials and methods

Sample

Subjects with ASD and their families were recruited at the Departments of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy at JW Goethe University Frankfurt/Main and at Saarland University Hospital as described previously.^{28,29} All parents and children had given informed consent. The study was approved by the local ethical committees (decision 162/99 (Frankfurt); 73/04 (Homburg), 237/09 (Frankfurt)). All patients were diagnosed using the gold-standard diagnostic tools Autism-Diagnostic Interview-Revised (ADI-R³⁰) and/or Autism Diagnostic Observation Schedule (ADOS-G³¹). Affected individuals met DSM-IV TR criteria for autism, Asperger's syndrome or PDD-NOS. We excluded any individuals with IQ<35, birth weight <1000g, cerebral palsy, reported chronic medical conditions (with the exception of epilepsy), blindness, deafness, known karyotypic or cytogenetic aberrations. The cohort in total consists of 510 male and 82 female patients from 592 families (492 trios, 73 duos, 27 singletons). 87.5% of index patients were diagnosed with strict autism including high- (IQ≥70) and low-functioning (IQ<70) autism. 12.5% received a spectrum diagnosis (Asperger's Syndrome or PDD-NOS). Of this cohort, we randomly selected 236 families as detection sample. The remaining set was used for

replication. Both samples were matched for diagnoses, age at diagnosis, sex and IQ (Supplementary Table 1).

Sequencing (detection sample)

Genomic DNA was extracted from blood or saliva using standard methods. For sequencing of the *CNTNAP2* promoter, PCR fragments were amplified covering the region from +585 to -709bp (NM_014141.5). This region spans translational start site, and known 5' transcription factor binding sites (data from ENCODE, UCSC Genome Browser, for full details on primers and PCR conditions, see Supplementary Table 2). Amplicons were purified using GeneJET™ PCR Purification Kit (Thermo Scientific, St. Leon-Rot, Germany). Sanger-sequencing of the detection sample was outsourced to SRD (Bad Homburg, Germany).

Genotyping (replication sample only)

Variants rs150447075, rs34712024 and rs71781329 were genotyped using restriction fragment length polymorphisms (RFLP; Watcut Database; Supplementary Table 2). A subset of N=295 families was analyzed using Custom TaqMan® SNP Assays (Life Technologies) AH5IQWH (rs150447075) and AH39SP9 (rs34712024), respectively. Details on setups are available upon request. In total we genotyped 99.73% (rs150447075 and rs34712024) and 98.37% (rs71781329) of the replication sample.

Prediction of transcription factor binding sites

To predict modified transcription factors binding sites (TFBS), Genomatix software MatInspector V2.6³² was used. Sequences including ± 50 bp of each variant were analyzed. Transcription factors were not considered for further investigation if bound to the reverse strand, or if standard score or matrix similarity were below 0.75, respectively.

Cloning

For construction of the luciferase vectors, a fragment of 1236 bp (+3-bp to -1239; NM_014141.5) was amplified from subjects carrying major alleles only or the respective minor alleles (Supplementary Table 2). This region spans the transcriptional start site and additional 500bp 5' of the predicted TF binding region. Amplicons were subcloned into pGEM®-T Easy Vector-System I before cloning into target vector pGL4.10[*luc2*] (Promega, Mannheim, Germany). Constructs carrying minor alleles of rs150447075, rs34712024, rs71781329 or major alleles were verified by direct sequencing.

Cell culture

HEK293T and SH-SY5Y cells were seeded at a density of 2×10^5 cells/cm² and grown in DMEM+GlutaMAX medium (all media from Life Technologies, Darmstadt, Germany, if not otherwise specified) supplemented with 10% FCS, 1% Pyruvate, 100U/ml Penicillin (PAA, Piscataway, NJ, USA), and 100µg/ml Streptomycin (PAA) in a humidified atmosphere at 37°C and 5% CO₂. For the differentiation assay, SH-SY5Y cells were seeded at a density of 5×10^3 cells/cm² in standard medium for 24h followed by treatment for additional 10 days with differentiation medium containing 10µM retinoic-acid (Sigma-Aldrich, St. Louis, MO, USA), 50ng/µl BDNF (Immunotools, Friesoythe, Germany), 2mM cAMP (Sigma-Aldrich), 20mM KCl (Sigma-Aldrich), 1XB27 supplement and 1XGlutamax in Neurobasal medium exchanging the medium every other day. Cells were harvested 24h after exchanging the media. Transition from undifferentiated to differentiated stages was confirmed via neuronal marker gene *MAPT* and cell division marker gene *CDC2* expression status (data not shown).

Relative quantitative RT real-time PCR

mRNA was purified using the GeneJet® mRNA-Purification Kit (Fermentas) and cDNA was generated with the Revert-Aid H-Plus Kit (Fermentas) according to the manufacturer's protocols. mRNA levels were analyzed adapting the UPL-System from Roche using the ABgene-ROX-Mastermix (Thermo Scientific Waltham, MA, USA) on a Step-one-plus™

system (Life Technologies). For details, see Supplementary Table 2. All experiments were controlled for PCR efficiency. Fold-change of mRNA expression level was calculated using the 2^{-ddCT} method.³³ *GAPDH* and *POLR2F* were used as stable reference genes as proposed by previous studies³⁴ and as confirmed in own experiments (data not shown).

Transfection and Luciferase Assay

Transfection of the luciferase vectors was performed using MetafectenePro (Biontex Laboratories; Planegg, Germany) according to the manufacturer's recommendations. Cells were plated at a density of 2×10^5 cells/cm² (mitotic cells) or 2.5×10^4 cells/cm² (differentiating SH-SY5Y) in a 12 well plate. Luciferase activity was measured by pooling two wells using the Dual-Luciferase[®] Reporter Assay System (Promega) following the manufacturer's protocol. Firefly luciferase activity was normalized to the empty vector pGL4.10[*luc2*]. All experiments were corrected for transfection efficiency based on co-transfection with Renilla luciferase expressing vector pGL4.74[*hRluc/TK*] (Promega). The experiment was performed in biological triplicate.

Statistical Analysis

Statistical analyses were done by IBM SPSS v20.0 or SAS v9.3. Power analysis was performed by Quanto v1.2.4³⁵ and G*Power v3.1.0 (<http://www.pscho.uni-duesseldorf.de/aap/projects/gpower>). Detection of Mendelian errors was performed using Haploview v4.1.³⁶ UNPHASED v3.1.4³⁷ was used to study single marker association, and to explore imprinting and sex-specific effects. As the ADI-R language items A9 (age at first word) and A10 (age at first phrase) were not normally distributed (Kolmogorov-Smirnov $p < 0.001$), non-parametric regression models to investigate the effect of the dichotomized genotypes (dominant model) with adjustment for age at diagnosis and IQ effects was performed by the SAS macro "npar" (<http://www.ams.med.uni->

goettingen.de/Projekte/makros/index.html). Luciferase assays were analysed by repeated measures analysis of variance (rANOVA).

Results

Promoter Sequencing

To identify known and novel variants, we sequenced the *CNTNAP2* promoter region 667 members of 236 German families with one child affected with ASD. Of the ten currently annotated SNPs (based on dbSNP, HapMap and UCSC data), we observed only three in our sample: rs150447075 (NG_007092.2:g.4595T>G), rs34712024 (NG_007092.2:g.4641A>G), and rs71781329 (NG_007092.2:g.4864_4865insGCGGCG) (Table 1). The STR rs71781329 is a six-fold GCG repeat (major allele termed here GCG[6]) where the insertion leads to eight repeats (GCG[8]). Interestingly, we also identified a previously unknown GCG[7] allele for the STR rs71781329. Furthermore, we identified six novel variants (M1-M6) previously not described in single individuals. As all genotype frequencies of the newly detected variants were below 1% and all variants were inherited, none of these variants was investigated further. The three annotated variants were selected for further single marker-based analyses (Table 1; Figure 1). No co-occurrence of any minor alleles was observed with the exception of one index patient and an unrelated mother carrying both minor alleles of rs150447075 and rs34712024. Thus, no haplotype-based association analysis was performed.

Single Marker Association Study in ASD

A preliminary Transmission Disequilibrium Test on the three variants in the detection sample showed a nominal association with ASD for rs34712024G (OR=0.20; CI95=0.04-0.91; p-value=0.016). For rs71781329, we observed complete transmission to the affected children (OR=∞; CI95=∞; p-value=0.041). We did not find an association for rs150447075G

(OR=1.00; CI95=0.32-3.10; p-value=1) (Table 2). Power analysis resulted in a power of 75% to replicate the detected OR of 0.20 for rs34712024G in our replication sample of 356 trios (p-value=0.05 two-sided, dominant model; minor allele frequency/MAF=0.017, population risk 1%). Likely due to low power, we did not identify any association in the replication sample and thus combined both sets for final analysis. The combined sample of 592 families had a power of 93% to detect an OR=0.2 (MAF 0.017, dominant model). In this full sample, again a nominal association for rs34712024G (OR=0.41; CI95=0.19-0.89; p-value=0.018) was observed (Table 2). For the trimeric STR rs71781329 (7+8 repeats combined), no association was observed in the full sample (OR=0.88; CI95=0.32-2.42; p-value=0.796). Descriptively, rs71781329GCG[7] was under-transmitted (OR=0.60; CI95=0.14-2.51; p-value=0.477), and rs71781329GCG[8] (OR=1.34; CI95=0.29-6.14; p-value=0.704) was over-transmitted when compared to rs71781329GCG[6] (Table 2). No gender-specific or imprinting effects were observed (data not shown).

Influence of promoter variants on ADI-R derived language measures

To explore if any of the three variants had an effect on language development non-parametric regression models were calculated. Age at diagnosis and IQ are known to influence language development and were included as covariates. We identified a suggestive effect (p-value=0.086) of the ASD-associated rs34712024G on age at first words (ADI-R, item A9) but not on age at first phrases (ADI-R, item A10, Table 3), with a slightly lower p-value in the male only subsample (p-value=0.067, male minor allele carriers ADI-R A9 mean = 34.8 months, SD=18.5, N=10; male major allele carriers ADI-R A9 mean= 24.4 months; SD=13.8; N=321; Table 3). Subjects with genotypes rs71781329GCG[6]/[7] but not GCG[6]/[8] showed a significantly (p-value < 0.00001) older age at first words and at first spoken phrase compared to GCG[6]/[6] (For details on mean differences see Table 3). However the number of subjects carrying the minor alleles was very low (GCG[6]/[7] N=2 and GCG[6]/[8] N=3); thus, these results have to be interpreted with caution.

Functional analyses of CNTNAP2 promoter variants

In silico analysis showed that all three promoter variants alter transcription factor binding sites (TFBS; Figure 1 and Supplementary Table 3): Carriers of the minor (G) allele of rs150447075 were predicted to carry an additional TFBS for the oncogenic transcription factor GLI3. The minor allele rs34712024G deletes TFBS of the TFs ZF300 and MYOG, and generates an additional recognition site for the kruppel-like ZF9. For rs7178329, an increased number of TFBS for EGR1 was predicted by the number of repeats (i.e. 3 TFBS for GCG[6], 4 TFBS for GCG[7], and 5 TFBS for GCG[8]).

In the differentiation assay, mRNA-expression levels of *CNTNAP2* were -3.77 fold down-regulated 72h post-induction of differentiation (PD) when compared to time-point 0h PD with a subsequent re-induction of expression reaching its peak at 216h PD (Figure 2A). The analysis of promoter activity at the critical time points PD (i.e. 0h, 72h and 216h) showed low activity for all variants at time points 0h and 72h PD and a remarkable increase at time point 216h (Figure 2B). This is in line with the observed re-induction of expression. No effect of the analyzed variants was observed at the level of overall activity regulation (all paired t-test p-values >0.3, data not shown). Subsequent analysis of the promoter activity in SH-SY5Y cells at each individual time point PD and in undifferentiated HEK293T cells showed cell line-specific and PD stage-dependent effects (Figure 2C and Supplementary Table 4): Compared to the major allele promoter, we observed a significantly increased activation of the rs150447075G promoter in undifferentiated SH-SY5Y (fold-change FC: 1.07, p-value=0.005) and HEK293T (FC=1.52, p-value=0.001) cells but not during neuronal differentiation. The ASD-associated rs34712024G allele significantly increased transcriptional efficiency in HEK293T cells (FC=1.61, p-value=0.006) but not at any differentiation stage of SH-SY5Y cells. The transcriptional efficiency of the rs71781329 minor alleles depended on the number of GCG repeats in HEK293T cells: GCG[7] showed a trend (p-value=0.055) towards an increased efficiency (1.219 fold;) whereas GCG[8] induced a

stronger and significant (p-value=0.022) 1.675 fold increase. However, this was not observed in the SH-SY5Y cell line. GCG[7] suggestively down-regulated efficiency in undifferentiated SH-SY5Y cells (FC=-1.03, p-value=0.061) only, whereas GCG[8] had a reducing effect at stage 72h PD (FC=-1.10, p-value=0.028), emphasizing the different nature of the two alleles as observed in the association analyses. No analyzed variant showed a significant effect on promoter activity at stage 216h PD.

Discussion

CNTNAP2 promoter variant rs37412024 is associated with ASD and may modulate language development

Our major finding of this study is a protective effect of the functional minor G-allele of rs34712024 on ASD. This allele increases transcriptional activity as shown in HEK293T cells. Testing three promoter variants of *CNTNAP2* for association with ASD showed a less frequent transmission of the minor G-allele of rs34712024 from parents to patients in the detection and the full sample. In addition, especially male carriers of the minor allele showed a trend for later first word acquisition assessed by retrospective parent report compared to the homozygous major allele carriers. This implies that the minor allele rs34712024G may also modulate language development. Previous studies also reported association of other, intronic *CNTNAP2* variants with language development, which implies *CNTNAP2* as a gene involved in language acquisition.^{10,11,13}

The protective role of rs34712024G on ASD risk is in contrast with the putative impact on language delay as measured by the ADI-R. Interestingly, contrasting effects of major and minor alleles of other *CNTNAP2* variants on the same phenotype but related to different disorders were described: The impact of rs2710102 on impaired non-word repetition test performance showed contrasting effects in children with specific language impairment (C allele as risk factor)¹¹ and dyslexia (T allele as risk factor).³⁸ If our bidirectional findings for rs34712024G are replicated by different groups and in different samples, this may point

towards pleiotropic effects of this allele, which may exert a modulating role in interaction with the remaining genetic background, and possibly also with environmental risk factors.

For STR rs71781329, we observed complete transmission of the minor alleles in the detection set. In the combined sample, we detected under-transmission of GCG[7] and over-transmission of GCG[8] compared to GCG[6], which did not reach significance, mainly due to the low frequency of these minor alleles. GCG[7] but not GCG[8] was associated with language delay. This finding, however, has to be interpreted with caution as only two individuals carrying the respective allele were observed in our sample.

Promoter variants influence gene expression of CNTNAP2

In our study, we took a closer look at the function of *CNTNAP2* promoter variants during neuronal differentiation. In silico analysis predicted that all three variants affect TFBS and indeed, our in vitro assay showed that all variants differentially regulated promoter activity as a function of cellular background and neuronal differentiation stage. Given that these variants alter TFBS, the protective effect of the minor allele rs34712024G on ASD may be attributed to altered binding of the two affected brain-specific TFs: ZF9, which is a strong enhancer for transcription during differentiation and cellular growth⁵ and TFAP4, a TF related to gene transcription upon DNA damage.³⁵ In parallel, the minor alleles GCG[7] and GCG[8] of rs71781329 may also act as modulators of transcription by inducing additional binding sites for EGR1, which plays a central role in the regulation of neuronal plasticity and differentiation. Synaptic activity, a mechanism putatively impaired in ASD³⁹, regulates EGR1 expression leading to altered transcription of targeted genes including *CNTNAP2*.⁴⁰ Interestingly, in the SH-SY5Y cell line model for neuronal differentiation, the GCG[7] repeat but not the GCG[8] repeat showed a suggestive reduction of the promoter efficiency in the pre-differentiating cells. During later stages of differentiation the efficiency is similar again. Furthermore, GCG[8] but not GCG[7] induced down regulation 72h after induction of differentiation, i.e. when *CNTNAP2* mRNA expression is lowest. Thus, differential *CNTNAP2*

expression in early stages of neuronal differentiation driven by the GCG[7] allele, but not in a later stage as observed for GCG[8] seems to modulate ASD language phenotypes.

The G allele of the non-associated rs150447075 showed an increase in transcriptional efficiency in mitotic cells (HEK293T and SH-SY5Y). GLI3, the putatively affected TF, is a non-brain-specific oncogene involved in limb formation. This fits very well with the results of the present association study, which did not show any effect of rs150447075 on ASD risk or development of language, whereas variants predicted to change TFBS relevant for neuronal regulation or response were (putative) protective or represent risk factors for ASD and delayed acquisition of language abilities.

CNTNAP2 expression may act as a bidirectional modulator of ASD and language development

In the brain *CNTNAP2* is ubiquitously expressed pre- and perinatally, but its expression is completely down-regulated during early childhood (age 1-6 years).⁴¹ Subsequently, *CNTNAP2* is re-expressed in all brain regions besides the striatum and the cerebellar cortex during life. This is in line with the reported down-regulation during neuronal differentiation of human neuronal precursor cells,⁴² and with our observation in the SH-SY5Y cell line model on neuronal differentiation. An aberrant transcriptional fine tuning or insufficient down-regulation of *CNTNAP2*, as observed here, may underlie the reported ASD-specific effects on language acquisition or hierarchical learning which are related to the striatum or the cerebellar cortex.^{43–46}

Conclusion

In summary, we performed a genetic and functional characterization study of the *CNTNAP2* promoter in ASD and report a bidirectional effect of rs34712024G on ASD, which might be mediated by an altered transcriptional efficiency of the *CNTNAP2* promoter. Our findings

need to be replicated in an independent and adequately powered ASD cohort. In addition, the role of rs71781329 should be studied in more detail in a larger ASD sample. The possible risk effects of rs34712024G and rs71781329GCG[7] on language development also need to be replicated in sufficiently powered and phenotypically well characterized samples, preferably in ASD, SLI, and typically developing children. The clarification of pleiotropic and multidirectional effects of ASD risk alleles as observed in this and other studies has to be a main target in future research to understand the functional impact of genetic risk factors and their interaction in ASD.

Acknowledgements

We thank all the families and patients for their cooperation and the clinical staff for their support in collecting data. We thank Heiko Zerlaut for database management, and Cornelia Wirth and Silvia Lindlar for excellent technical assistance.

Conflict of Interest

None of the authors has to report any conflict of interest. Talks of Christine Freitag have been supported by Shire, Novartis, and Eli Lilly; and she has acted as consultant for Desitin.

Supplementary Information

Supplementary information is available at Molecular Psychiatry's website.

References

- 1 Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D *et al.* Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* 2006; **368**: 210–215.
- 2 Brugha TS, McManus S, Bankart J, Scott F, Purdon S, Smith J *et al.* Epidemiology of autism spectrum disorders in adults in the community in England. *Arch Gen Psychiatry* 2011; **68**: 459–465.
- 3 World Health Organization *International classification of mental and behavioral disorders. Clinical descriptions and diagnostic guidelines*. World Health Organization: Geneva, 1992.
- 4 American Psychiatric Association *Diagnostic and statistical manual of mental disorders: DSM-IV-TR, 4th edn*. American Psychiatric Association: Washington, DC, 2000.
- 5 Lichtenstein P, Carlström E, Råstam M, Gillberg C, Anckarsäter H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am J Psychiatry* 2010; **167**: 1357–1363.
- 6 Freitag CM, Staal W, Klauck SM, Duketis E, Waltes R. Genetics of autistic disorders: review and clinical implications. *Eur Child Adolesc Psychiatry* 2010; **19**: 169–178.
- 7 Peñagarikano O, Geschwind DH. What does CNTNAP2 reveal about autism spectrum disorder? *Trends Mol Med* 2012; **18**: 156–163.
- 8 Alarcón M, Cantor RM, Liu J, Gilliam TC, Geschwind DH. Evidence for a Language Quantitative Trait Locus on Chromosome 7q in Multiplex Autism Families. *Am J Hum Genet* 2002; **70**: 60–71.
- 9 Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M *et al.* A Common Genetic Variant in the Neurexin Superfamily Member CNTNAP2 Increases Familial Risk of Autism. *Am J Hum Genet* 2008; **82**: 160–164.

- 10 Alarcón M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM *et al.* Linkage, Association, and Gene-Expression Analyses Identify CNTNAP2 as an Autism-Susceptibility Gene. *Am J Hum Genet* 2008; **82**: 150–159.
- 11 Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M *et al.* A Functional Genetic Link between Distinct Developmental Language Disorders. *N Engl J Med* 2008; **359**: 2337–2345.
- 12 Newbury DF, Paracchini S, Scerri TS, Winchester L, Addis L, Richardson AJ *et al.* Investigation of Dyslexia and SLI Risk Variants in Reading- and Language-Impaired Subjects. *Behav Genet* 2011; **41**: 90–104.
- 13 Whitehouse AJO, Bishop DVM, Ang QW, Pennell CE, Fisher SE. CNTNAP2 variants affect early language development in the general population. *Genes Brain Behav* 2011; **10**: 451–456.
- 14 Whalley HC, O'Connell G, Sussmann JE, Peel A, Stanfield AC, Hayiou-Thomas ME *et al.* Genetic variation in CNTNAP2 alters brain function during linguistic processing in healthy individuals. *Am J Med Genet* 2011; **156**: 941–948.
- 15 Eyler LT, Pierce K, Courchesne E. A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* 2012; **135**: 949–960.
- 16 Bakaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM *et al.* Molecular Cytogenetic Analysis and Resequencing of Contactin Associated Protein-Like 2 in Autism Spectrum Disorders. *Am J Hum Genet* 2008; **82**: 165–173.
- 17 O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 2011; **43**: 585–589.
- 18 Caselli R, Mencarelli MA, Papa FT, Ariani F, Longo I, Meloni I *et al.* Delineation of the phenotype associated with 7q36.1q36.2 deletion: long QT syndrome, renal hypoplasia and mental retardation. *Am J Med Genet A* 2008; **146**: 1195–1199.

- 19 Rossi E, Verri AP, Patricelli MG, Destefani V, Ricca I, Vetro A *et al.* A 12Mb deletion at 7q33–q35 associated with autism spectrum disorders and primary amenorrhea. *Eur J Med Genet* 2008; **51**: 631–638.
- 20 Petrin AL, Giacheti CM, Maximino LP, Abramides DVM, Zanchetta S, Rossi NF *et al.* Identification of a microdeletion at the 7q33–q35 disrupting the CNTNAP2 gene in a Brazilian stuttering case. *Am J Med Genet* 2010; **152**: 3164–3172.
- 21 Poot M, Beyer V, Schwaab I, Damatova N, Slot R, Prothero J *et al.* Disruption of CNTNAP2 and additional structural genome changes in a boy with speech delay and autism spectrum disorder. *Neurogenetics* 2010; **11**: 81–89.
- 22 Sehested LT, Møller RS, Bache I, Andersen NB, Ullmann R, Tommerup N *et al.* Deletion of 7q34–q36.2 in two siblings with mental retardation, language delay, primary amenorrhea, and dysmorphic features. *Am J Med Genet* 2010; **152**: 3115–3119.
- 23 Nord AS, Roeb W, Dickel DE, Walsh T, Kusenda M, O'Connor KL *et al.* Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. *Eur J Hum Genet* 2011; **19**: 727–731.
- 24 Peñagarikano O, Abrahams BS, Herman EI, Winden KD, Gdalyahu A, Dong H *et al.* Absence of CNTNAP2 Leads to Epilepsy, Neuronal Migration Abnormalities, and Core Autism-Related Deficits. *Cell* 2011; **147**: 235–246.
- 25 Poliak S. Juxtaparanodal clustering of Shaker-like K⁺ channels in myelinated axons depends on Caspr2 and TAG-1. *J Cell Biol* 2003; **162**: 1149–1160.
- 26 Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S *et al.* Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996; **274**: 1527–1531.
- 27 Shastri BS. SNPs: impact on gene function and phenotype. *Methods Mol Bio* 2009; **578**: 3–22.
- 28 Klauck SM, Poustka F, Benner A, Lesch KP, Poustka A. Serotonin transporter (5-HTT) gene variants associated with autism? *Hum Mol Genet* 1997; **6**: 2233–2238.

- 29 Freitag CM, Agelopoulos K, Huy E, Rothermundt M, Krakowitzky P, Meyer J *et al.* Adenosine A(2A) receptor gene (ADORA2A) variants may increase autistic symptoms and anxiety in autism spectrum disorder. *Eur Child Adolesc Psychiatry* 2010; **19**: 67–74.
- 30 Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 1994; **24**: 659–685.
- 31 Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC *et al.* The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 2000; **30**: 205–223.
- 32 Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 1995; **23**: 4878–4884.
- 33 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001; **25**: 402–408.
- 34 Leduc V, Legault V, Dea D, Poirier J. Normalization of gene expression using SYBR green qPCR: a case for paraoxonase 1 and 2 in Alzheimer's disease brains. *J. Neurosci Methods* 2011; **200**: 14–19.
- 35 Gauderman W MJ. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. <http://hydra.usc.edu/gxe>, 2006
- 36 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
- 37 Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008; **66**: 87–98.
- 38 Peter B, Raskind WH, Matsushita M, Lisowski M, Vu T, Berninger VW *et al.* Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *J Neurodevelop Disord* 2011; **3**: 39–49.

- 39 Ebert DH, Greenberg ME. Activity-dependent neuronal signalling and autism spectrum disorder. *Nature* 2013; **493**: 327–337.
- 40 Knapska E, Kaczmarek L. A gene for neuronal plasticity in the mammalian brain: Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK? *Prog Neurobiol* 2004; **74**: 183–211.
- 41 Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M *et al.* Spatio-temporal transcriptome of the human brain. *Nature* 2011; **478**: 483–489.
- 42 Konopka G, Wexler E, Rosen E, Mukamel Z, Osborn GE, Chen L *et al.* Modeling the functional genomics of autism using human neurons. *Mol Psychiatry* 2012; **17**: 202–214.
- 43 Haas RH, Townsend J, Courchesne E, Lincoln AJ, Schreibman L, Yeung-Courchesne R. Neurologic abnormalities in infantile autism. *J. Child Neurol* 1996; **11**: 84–92.
- 44 Langen M, Schnack HG, Nederveen H, Bos D, Lahuis BE, Jonge MV de *et al.* Changes in the developmental trajectories of striatum in autism. *Biol. Psychiatry* 2009; **66**: 327–333.
- 45 Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. *Annu Rev Neurosci* 2009; **32**: 413–434.
- 46 Desrochers TM, Badre D. Finding parallels in fronto-striatal organization. *Trends Cogn Sci* 2012; **16**: 407–408.

Table 1 Variants detected in the sequenced sample

	SNP	Alleles major/minor	N Index (236)	Inheritance	MAF (index only) ^a	Position (GRCh37)
Annotated	rs150447075	T/G	11	3 pat.; 5 mat.; 3 NA	G=0.023	145813047
	rs34712024	A/G	5	1 pat.; 1 mat.; 3 NA	G=0.011	145813093
	rs71781329	GCG[6]/GCG[7]	1	1 pat.	GCG[7]=0.002	145813308
		GCG[6]/GCG[8]	2	1 pat.; 1 mat.	GCG[8]=0.004	
	combined	GCG[6]/GCG(7_8)	3	2 pat.; 1 mat.	GCG(7_8)=0.006	
	rs13225855		ND			145813233
	rs13225866		ND			145813252
	rs13225872		ND			145813261
	rs13226652		ND			145813270
	rs13225963		ND			145813304
	rs13225979		ND			145813374
	rs13226843		ND			145813386
Novel	M1	–/C		Detected in half sibling only		145813046
	M2	A/G	1	1 pat.		145813238
	M3	C/G		Detected in mother only		145813410
	M4	C/G	1	1 pat.		145813427
	M5	A/G	1	1 pat.		145813497
	M6	G/C	1	1 mat.		145813559

Abbreviations: GRCh37, Genome Reference Consortium human genome (build 37); MAF, minor allele frequency; mat., maternal; NA, not applicable (parents not available); ND, not detected; pat., paternal
a) Minor allele frequency within our detection sample.

Table 2 Single marker transmission disequilibrium test

SNP (minor/major alleles)	MAF^a	OR minor allele	p-value^b
Detection sample (N=236 families)			
rs150447075T/G	0.020	1.00, CI95: 0.32-3.10	1.000
rs34712024A/G	0.017	0.20, CI95: 0.04-0.91	0.016*
rs71781329GCG[6]/GCG(7_8)	0.005	∞, CI95: ∞ (complete transmission)	0.041*
Replication sample (N=356 families)			
rs150447075T/G	0.025	0.89, CI95: 0.45-1.74	0.732
rs34712024A/G	0.015	0.58, CI95: 0.23-1.48	0.249
rs71781329GCG[6]/GCG(7_8)	0.006	0.49, CI95: 0.14-1.73	0.239
Combined sample (N=592 families)			
rs150447075T/G	0.023	0.92, CI95: 0.51-1.64	0.768
rs34712024A/G	0.016	0.41, CI95: 0.19-0.89	0.018*
rs71781329GCG[6]/GCG(7_8)	0.006	0.88, CI95: 0.32-2.42	0.796
rs71781329GCG[6]/GCG[7] ^c	0.003	0.60, CI95: 0.14-2.51	0.477
rs71781329GCG[6]/GCG[8] ^d	0.003	1.34, CI95: 0.29-6.14	0.704

Abbreviations: CI, confidence interval; MAF, minor allele frequency (calculated on all samples); OR, odds ratio.

a) Minor allele frequency within our ASD sample.

b) p-values <0.05 are marked by an asterisk.

c) Families of rs71781329GCG[8] allele carriers were omitted; N=588 families were included into analysis.

d) Families of rs71781329GCG[7] allele carriers were omitted; N=584 families were included into analysis.

Table 3 Non-parametric regression models for genetic effects on ADI-R language measures

Model	<i>first words</i> (ADI-R A9)			<i>first phrase</i> (ADI-R A10)			<i>first words - male only</i> (ADI-R A9)			<i>first phrase - male only</i> (ADI-R A10)		
	T	p-value	T	p-value	T	p-value	T	p-value	T	p-value	T	p-value
rs150447075 AA vs AG/GG	0.979	0.322	0.943	0.331	0.265	0.606	0.265	0.606	0.175	0.676		
Age	-0.290	0.772	0.184	0.854	-0.154	0.877	-0.154	0.877	0.385	0.700		
IQ	-4.942	<0.0001	-6.039	<0.0001	-4.424	<0.0001	-4.424	<0.0001	-5.652	<0.0001		
N (carriers G)	383(17)		336(16)		331(14)		331(14)		292(13)			
Mean (SD)	TT: 25.48 (14.42) TG/GG: 21.65 (9.51)		TT: 37.69 (17.95) TG/GG: 33.94 (10.47)		TT: 24.82 (14.14) TG/GG: 21.93 (10.41)		TT: 24.82 (14.14) TG/GG: 21.93 (10.41)		TT: 37.01 (18.05) TG/GG: 33.77 (10.19)			
rs34712024 TT vs TG/GG	2.953	0.086	0.016	0.899	3.362	0.067	3.362	0.067	0.027	0.871		
Age	-0.277	0.781	0.176	0.861	-0.142	0.887	-0.142	0.887	0.375	0.707		
IQ	-4.808	<0.0001	-5.968	<0.0001	-4.293	<0.0001	-4.293	<0.0001	-5.602	<0.0001		
N (carriers G)	383(10)		336(8)		331(10)		331(10)		292(8)			
Mean (SD)	AA: 25.05 (14.07) AG/GG: 34.80 (18.50)		AA: 37.36 (17.52) AG/GG: 43.63 (23.80)		AA: 24.39 (13.76) AG/GG: 34.80 (18.50)		AA: 24.39 (13.76) AG/GG: 34.80 (18.50)		AA: 36.67 (17.60) AG/GG: 43.63 (23.80)			
rs71781329 GCG[6]/[6] vs [6]/(7_8)	2.689	0.101	0.581	0.446	1.207	0.272	1.207	0.272	0.039	0.843		
Age	-0.180	0.858	0.237	0.813	-0.005	0.996	-0.005	0.996	0.447	0.655		
IQ	-4.866	<0.0001	-6.036	<0.0001	-4.329	<0.0001	-4.329	<0.0001	-5.664	<0.0001		
N (carriers GCG(7_8))	380(5)		334(8)		328(4)		328(4)		290(4)			
Mean (SD)	GCG[6]/[6]: 25.03 (13.62) GCG[6]/(7_8): 45.00 (37.11)		GCG[6]/[6]: 37.22 (16.98) GCG[6]/(7_8): 58.80 (42.51)		GCG[6]/[6]: 24.44 (13.27) GCG[6]/(7_8): 44.25 (42.81)		GCG[6]/[6]: 24.44 (13.27) GCG[6]/(7_8): 44.25 (42.81)		GCG[6]/[6]: 36.63 (17.07) GCG[6]/(7_8): 55.50 (48.34)			
rs71781329 GCG[6]/[6] vs [6]/[7]	83.972	<0.0001	19.135	<0.0001								
Age	-0.304	0.761	0.084	0.9341								
IQ	-4.954	<0.0001	-6.249	<0.0001								
N (carriers GCG[7])	377(2)		331(2)									
Mean (SD)	GCG[6]/[6]: 25.03 (13.62) GCG[6]/[7]: 78.00 (42.43)		GCG[6]/[6]: 37.22 (16.98) GCG[6]/[7]: 99.00 (38.18)		Not performed as all carriers were male							
rs71781329 GCG[6]/[6] vs [6]/[8]	0.109	0.741	0.104	0.748								
Age	-0.309	0.757	0.046	0.964								
IQ	-4.857	<0.0001	-6.078	<0.0001								
N (carriers GCG[8] allele)	378(3)		332(3)									
Mean (SD)	GCG[6]/[6]: 25.03 (13.62) GCG[6]/[8]: 23.00 (6.25)		GCG[6]/[6]: 37.22 (13.98) GCG[6]/[8]: 32.00 (13.86)		Not performed as all carriers were male							

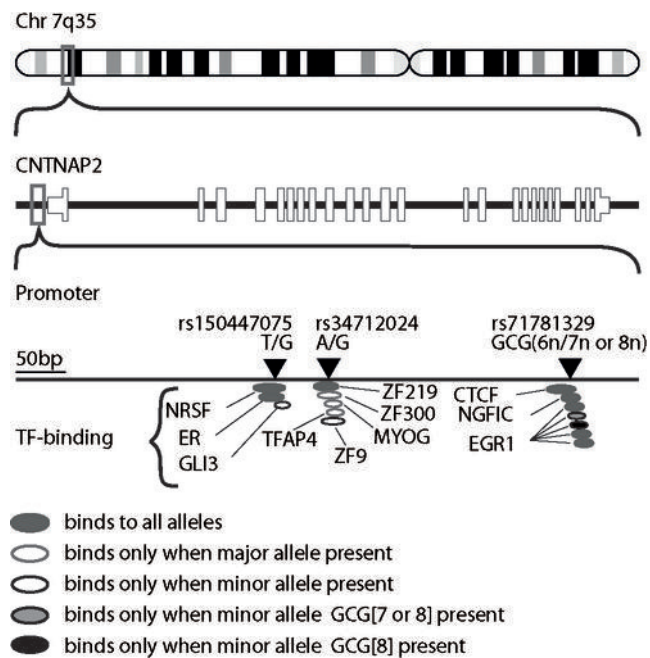
N, Total number of samples included in the model with carriers of the respective minor alleles in parentheses; vs, versus; T, T-value of non-parametric regression model; Mean, mean months at first words or first phrase, respectively, SD, standard deviation.

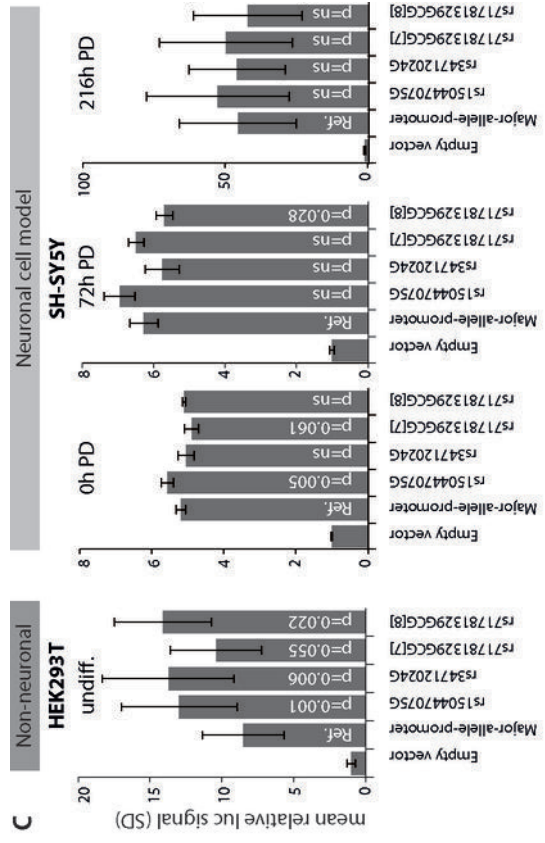
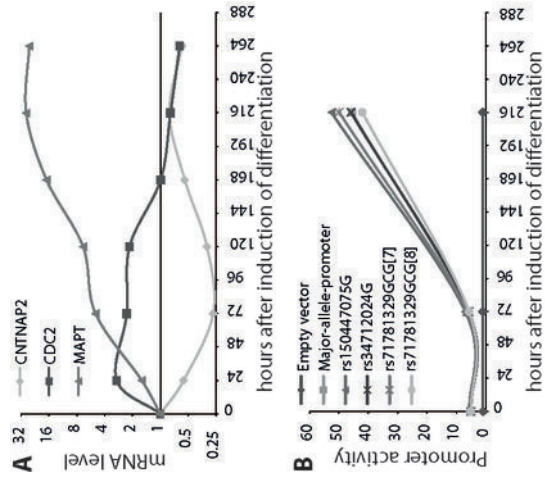
Figure 1: Variants of the *CNTNAP2* promoter. Variants analyzed in this study, their schematic localization within the respective region and the predicted transcription factor binding sites (TFBS) in proximity to the different minor alleles are shown. TFs not binding to the variants or interacting with the reverse strand only are omitted. Approximate length and localization of binding sites are depicted by size and position of the ellipses, respectively. For details on the corresponding prediction matrices and anchor sequences see Supplementary Table 3.

Abbreviations: CTCF, Insulator protein CTCF (CCCTC-binding factor); EGR1, early growth response 1; ER, Estrogen response element; GLI3, GLI-Kruppel family member 3; MYOG, Myogenic bHLH protein myogenin (myf4); NGFIC, Nerve growth factor-induced protein C; NRSF, Neuron-restrictive silencer factor; TFAP4, Activator protein 4; ZF9, Core promoter-binding protein (CPBP) with 3 Kruppel-type zinc fingers (also known as KLF6); ZF219, Kruppel-like zinc finger protein 219; ZF300, KRAB-containing zinc finger protein 300.

Figure 2: Functional analysis of promoter activation dependent on cell type, SH-SY5Y differentiation status and genetic variants. **A:** mRNA expression of *CNTNAP2* during neuronal differentiation of SH-SY5Y. *CNTNAP2* is down-regulated until 72h after induction of differentiation and re-induced afterwards. Neuronal marker *MAPT* is up-regulated upon induction of differentiation, whereas cell division marker *CDC2* shows down-regulation after 24h post-induction. **B:** Luciferase reporter gene assay shows a constant low promoter activity in early stages of differentiation (0h-72h) and a marked increase 216h after induction. **C:** Promoter activity is influenced by genetic variants depending on cell type and differentiation status. Note that the expression efficiency is higher in the HEK293T cells than in SH-SY5Y undifferentiated cells and highest in fully differentiated SH-SY5Y cells. P-values show significant differences between promoter with minor alleles compared to the promoter with major alleles only. All assays were normalized to the luciferase reporter construct without any promoter (empty vector).

PD, post-induction of differentiation; SD, standard deviation.





Supplementary information for

Promoter variants in the ASD risk-gene *CNTNAP2*: a genetic and functional approach

Marnie Kopp, MSc¹, Andreas Chiocchetti, PhD¹, Regina Waltes, PhD¹, Eftichia Duketis, MD¹, Tomasz Jarczok, MD¹, Fritz Poustka, MD¹, Anette Voran, MD², Ulrike Graab, MSc³, Jobst Meyer, PhD⁴, Sabine M. Klauck, PhD⁵, Simone Fulda, MD³, Christine M. Freitag, MD^{1*}

¹ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, JW Goethe University, Frankfurt am Main, Germany

² Department of Child and Adolescent Psychiatry, Saarland University, Homburg, Germany

³ Institute of Experimental Cancer Research in Pediatrics, Frankfurt am Main, Germany

⁴ Department of Neurobehavioral Genetics, Institute of Psychobiology, University of Trier, Germany

⁵ Division of Molecular Genome Analysis, German Cancer Research Center (DKFZ), Heidelberg, Germany

* Corresponding author

Professor Dr. med. Dipl. theol. Christine M. Freitag, MD, MA

Supplementary Table 1

Descriptive statistics of the detection, the validation and the combined sample _____ 2

Supplementary Table 2

Detailed PCR conditions for genotyping of the variants of interest _____ 3

Supplementary Table 3

Transcription factors predicted to bind proximal to analyzed variants _____ 4

Supplementary Table 4

Relative luciferase activity comparing minor to major alleles _____ 5

Supplementary information

Databases and websites used in this study _____ 6

Supplementary Table 1

Descriptive statistics of the detection, the validation and the combined sample

	N	Age in years (SD)	N	IQ (SD)	N	Sex (frequency)	N	Diagnosis (frequency)
Detection	216	10.35 (5.44)	207	89.13 (25.74)	236	M (0.869) F (0.131)	236	A (0.856) S (0.144)
Validation	353	11.00 (6.48)	317	88.10 (23.18)	356	M (0.857) F (0.143)	356	A (0.888) S (0.112)
p-value		U-test: 0.725		U-Test: 0.407		Chi ² : 0.717		Chi ² : 0.256
Combined	569	10.75 (6.11)	524	88.51 (24.20)	592	M (0.861) F (0.139)	592	A (0.875) S (0.125)

Abbreviations: A, autism; F, female; M, male; S, spectrum; SD, standard deviation.

Supplementary Table 2 Primer and PCR conditions used in this study here

Sequencing		Primer (5' - 3')	Polymerase
Amplicon	+585bp to -709bp (NM_014141.5)	Fw: TGCATTAAACGGCATCTGCA Rv: TGCAGAGGCAGCTGCTGAC	TopTaq™ (Qiagen, Hilden, Germany); add Q- solution
Genotyping RFLP		Primer (5' - 3')	Polymerase
Variant	rs150447075T>G	Fw: GTGCATTAAACGGCATCTGCA Rv: CCCTGGCCTCGTTTGAAAG	DreamTaq™ (Thermo Scientific, St. Leon-Rot, Germany.)
	rs34712024A>G	Fw: GTGCATTAAACGGCATCTGCA Rv: CCCTGGCCTCGTTTGAAAG	DreamTaq™ (Thermo Scientific)
rs71781329 GCG[6]>GCG[7_8]		Fw: CAGGGCAGAAAGGGTTTGGACA Rv: CTCACATGGAGAGCTTGTGTATGTG	AmpliTaq® Gold + 0.8% DMSO (Life Technologies, Darmstadt, Germany)
Cloning		Primer (5' - 3')	Polymerase
Amplicon	Promoter insert +3-bp to -1239 (NM_014141.5)	Fw: g ^a cagcATTCCAGGATTAGGCAAGCCCTCCA Rv: a ^a gcctTTGTGTATGTGGGAGCGCGCA	PRECISOR High-Fidelity DNA Polymerase (BioCat, Heidelberg, Germany)
mRNA Expression		Primer (5' - 3')	UPL-probe number (Reaction volume)
Target	GAPDH	Fw: AGCCACATCGCTCAGACAC Rv: GCCCAATACGACCAATCC	#60 (20µl)
	POL2RF	Fw: GATGAAGGGCTAGATGACTTGG Rv: CCAGAGGGGAGGATCTCG	#2 (20µl)
CNTNAP2 NM_014141.5		Fw: CCTGGAGCTCTACAGCCAGT Rv: TCACCCATTCACTCTCAA	#26 (20µl)
	MAPT NM_016835	Fw: ACCACAGCCACCTTCTCCT Rv: CAGCCATCCTGGTTCAAAGT	#55 (20µl)
CDC2 NM_001786		Fw: TGGATCTGAAGAAATACTTGGATTCTA Rv: CAATCCCCTGTAGGATTGG	#79 (20µl)
		Rest. enzyme	PCR efficiency
		Cac8I (NEB)	96.03%
		AluI (NEB)	98.61%
		HpaII (NEB)	99.42%
			95.67%
			93.04%
		Fragment lengths	Comment
		T: 323, 85 and 1bp G: 307, 85, 16 and 1bp	Reference Gene
		A: 219, 133 and 57bp G: 276 and 133bp	Reference Gene
		GCG[6]: 139, 80, 49 and 45bp GCG[7]: 139, 83, 49 and 45bp GCG[8]: 139, 86, 49 and 45bp	Target gene
			Neuronal marker
			Cell division marker

Abbreviations: DMSO, dimethyl sulfoxide; Fw, forward; NEB, New England Biolabs, Frankfurt, Germany; Rest. Enzyme, restriction enzyme; Rv, reverse

Supplementary Table 3 Transcription factors predicted to bind proximal to analyzed variants^a

rs150447075T				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
NRSF	Neuron-restrictive silencer factor	1	0.97	acttcAGCAccgcggacagcatcacctccg
ER	Estrogen response elements, IR3 sites	0.809	0.861	caccgcggacagcatCACtccc
rs150447075G				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
NRSF	Neuron-restrictive silencer factor	1	0.97	acttcAGCAccgcggacagcatcaccgccc
ER	Estrogen response elements, IR3 sites	0.809	0.835	caccgcggacagcatCACgccc
GLI3 ^c	GLI-Kruppel family member GLI3	0.801	0.88	ccgcCCGCccaggga
rs34712024A				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
ZF219	Kruppel-like zinc finger protein 219	1	0.996	ccttgctcgcCCCCgccccagc
ZF300 ^d	KRAB-containing zinc finger protein 300	1	1	tgctccgccccgcCCCCagctga
MYOG ^d	Myogenic bHLH protein myogenin (myf4)	1	0.908	ccgcccCAGCtgatgga
TFAP4 ^d	Activator protein 4	1	0.949	cgccccAGCTgatggac
rs34712024G				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
ZF219	Kruppel-like zinc finger protein 219	1	0.996	ccttgctcgcCCCCgccccggc
ZF9 ^c	Core promoter-binding protein (CPBP) with 3 Kruppel-type zinc fingers (KLF6)	1	0.925	ttgtccgccCCGCccccgctg
rs71781329GCG[6]				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
CTCF	Insulator protein CTCF (CCCTC-binding factor)	0.863	0.796	ccgggctcacaggagcTG GCggcggcg
NGFIC	Nerve growth factor-induced protein C	0.786	0.812	gctg GCGGcgggcggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcgt
EGR1	Early growth response 1	0.789	0.894	ggcggcggCGGCgtctc
rs71781329GCG[7]				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
CTCF	Insulator protein CTCF	0.863	0.796	ccgggctcacaggagcTG GCggcggcg
NGFIC	Nerve growth factor-induced protein C	0.786	0.812	gctg GCGGcgggcggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcgt
EGR1 ^c	Early growth response 1	0.789	0.894	ggcggcggCGGCgtctc
rs71781329GCG[8]				
TF	Full name	Core sim.	Matrix sim.	Sequence ^b
CTCF	Insulator protein CTCF	0.863	0.796	ccgggctcacaggagcTG GCggcggcg
NGFIC	Nerve growth factor-induced protein C	0.786	0.812	gctg GCGGcgggcggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1	Early growth response 1	0.789	0.91	ggcggcggCGGCggcg
EGR1 ^c	Early growth response 1	0.789	0.91	ggcggcggCGGCggcgt
EGR1 ^c	Early growth response 1	0.789	0.894	ggcggcggCGGCgtctc

Abbreviations: sim., similarity; TF, transcription factor.

a) Corresponding prediction matrices and anchor sequences retrieved from MatInspector.

b) Predicted matrices of TF binding sites; anchor sequences are displayed in capital letters. Variants are depicted in bold.

c) Additional transcription factor binding site.

d) Deleted transcription factor binding site.

Supplementary Table 4 Relative luciferase activity of *CNTNAP2* promoter depending on genetic variants and cell types

HEK293T	Mean ^a	SD	FC	p-value ^b
Empty	1.00	0.28	NA	NA
Major alleles only	8.51	2.85	1.00	NA
rs150447075G	12.96	4.02	1.52	0.001
rs34712024G	13.72	4.58	1.61	0.006
rs71781329GCG[7]	10.40	3.18	1.22	0.055
rs71781329GCG[8]	14.11	3.36	1.66	0.022

SH-SY5Y 0h PD	Mean ^a	SD	FC	p-value ^b
Empty	1.00	0.02	NA	NA
Major alleles only	5.20	0.13	1.00	NA
rs150447075G	5.57	0.17	1.07	0.005
rs34712024G	5.05	0.22	-1.03	0.106
rs71781329GCG[7]	4.89	0.20	-1.06	0.061
rs71781329GCG[8]	5.11	0.04	-1.02	0.324

SH-SY5Y 72h PD	Mean ^a	SD	FC	p-value ^b
Empty	1.00	0.07	NA	NA
Major alleles only	6.27	0.40	1.00	NA
rs150447075G	6.95	0.42	1.11	0.231
rs34712024G	5.75	0.48	-1.09	0.414
rs71781329GCG[7]	6.49	0.21	1.03	0.571
rs71781329GCG[8]	5.69	0.23	-1.10	0.028

SH-SY5Y 216h PD	Mean ^a	SD	FC	p-value ^b
Empty	1.00	0.34	NA	NA
Major alleles only	45.60	20.47	1.00	NA
rs150447075G	52.59	25.14	1.15	0.491
rs34712024G	45.89	17.06	1.01	0.966
rs71781329GCG[7]	49.88	23.34	1.09	0.611
rs71781329GCG[8]	42.03	19.14	-1.09	0.544

Abbreviations: FC, symmetric fold change with negative values corresponding to down-regulations and positive values to upregulations of respective promoter constructs compared to major alleles only vector; PD, Post differentiation; SD, Standard deviation of three biological replicates

a) Mean: Luciferase activity of three biological replicates normalized to empty vector.

b) p-values were calculated using ANOVA for repeated measures comparing respective constructs to major alleles only vector.

Supplementary information Databases used in this study

dbSNP: Version 3.1, March 8, 2006; <http://www.ncbi.nlm.nih.gov/projects/SNP/>

ENCODE: June 2005 Freeze; <http://genome.ucsc.edu/ENCODE/>

Ensemble: Ensembl release 56 – September 2009; <http://www.ensembl.org/index.html>

Hapmap: Data Rel 27 Phase+III; <http://hapmap.ncbi.nlm.nih.gov/>

MatInspector: genomatrix software suite v2.6; <http://www.genomatix.de>

USCS: Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly; <http://genome.ucsc.edu/>

Watcut: June 19, 2007. Enzymes as of July 2006, Rebase; <http://watcut.uwaterloo.ca/>